

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



**Evaluating end-of-life care : do berieved relatives provide accurate information on patients' experience of pain, anxiety and depression?**

McPherson, Christine Jane

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

**END USER LICENCE AGREEMENT**



**Unless another licence is stated on the immediately following page** this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

**Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



**EVALUATING END-OF-LIFE CARE: DO BEREAVED  
RELATIVES PROVIDE ACCURATE INFORMATION ON  
PATIENTS' EXPERIENCES OF PAIN, ANXIETY AND  
DEPRESSION?**

By

Christine J. McPherson  
*R.N., B.Sc., M.Sc.*

A thesis submitted to King's College London  
for the degree of Doctor of Philosophy

GUY'S, KING'S AND ST. THOMAS' SCHOOL OF MEDICINE  
DEPARTMENT OF PALLIATIVE CARE AND POLICY  
WESTON EDUCATION CENTRE  
CUTCOMBE ROAD DENMARK HILL  
LONDON SE5 8RJ  
UNITED KINGDOM

July 2003

© Christine J. McPherson

# Abstract

Research into end-of-life care presents a major challenge because it is difficult to obtain the views and experiences of representative samples of patients. Studies relying on patients' accounts prior to death are potentially biased, as they only represent that proportion of patients with an identifiable terminal illness, who are able and willing to participate. An alternative approach that overcomes many of these problems is the retrospective or "after death" approach. Here, observations are gathered from proxies acting on behalf of the patient, after the patient's death. Usually this is a relative involved in their care. This approach has been used in a number of influential studies and has provided important insights into death and dying in the late 20th century. In spite of its widespread use, little is known about the validity of proxies' accounts. This thesis presents a systematic review of studies that have examined the validity of the approach by comparing proxies' accounts with those of the patient. Agreement was shown to be poorer for aspects of patients' experiences that were more subjective in nature, for example pain, anxiety and depression. Focusing on these aspects, a series of research questions were posed and investigated. These included proxies' comprehension of surveys questions, judgments regarding patients' symptoms, the influence of proxies' own emotions on their reports and consistency of responses over time. Using qualitative and quantitative methodologies thirty bereaved relatives were interviewed approximately three months after the patient's death, and followed-up 3-5 months later. These investigations provide researchers with a greater awareness of reasons why disagreement between patients and proxies accounts might occur. Understanding possible biases and errors will facilitate the development of methods to improve the validity of proxies' accounts. In addition to this, recommendations are made to improve existing measures and for future research.

# Acknowledgments

I would like to take this opportunity to thank all those individuals who assisted with the completion of this Ph.D. To my supervisor, Professor Julia Addington-Hall, I offer my gratitude for believing in me and for providing support and encouragement. Regarding my colleagues in the Department of Palliative Care and Policy, I thank them for their contributions. My appreciation also goes to the hospice service for providing practical assistance on matters relating to the recruitment and welfare of participants. Finally, I would like to acknowledge the generosity and fortitude of all of the individuals who volunteered to take part in the research, and shared their thoughts and feelings with me at such a difficult time.



*This work is dedicated to my husband, Douglas.*

# Table of Contents

|   |           |
|---|-----------|
| Abstract . . . . .  | 2         |
| Acknowledgments . . . . .   | 3         |
| Table of Contents. . . . .  | 5         |
| List of Figures . . . . .   | 9         |
| List of Tables. . . . .   | 10        |
| <b>CHAPTER 1: Introduction</b>  | <b>11</b> |
| 1.1 Methodological problems when evaluating care of the dying . . . . . | 12        |
| 1.2 Retrospective approach . . . . .                                    | 12        |
| 1.3 Proxies in other areas of health . . . . .                          | 13        |
| 1.4 Significant others as proxies . . . . .                             | 13        |
| 1.4.1 Epidemiology. . . . .   | 13        |
| 1.4.2 Quality of life. . . . .  | 15        |
| 1.4.3 Health care decisions . . . . .                                   | 19        |
| 1.5 Health professionals as proxies . . . . .                           | 19        |
| 1.5.1 Health care decisions . . . . .                                   | 20        |
| 1.5.2 Quality of life. . . . .  | 20        |
| 1.6 Discussion . . . . .  | 23        |
| <b>CHAPTER 2: Proxy systematic review</b>                               | <b>25</b> |
| 2.1 Introduction . . . . .  | 25        |
| 2.2 Method . . . . .  | 25        |
| 2.2.1 Literature search strategy . . . . .                              | 25        |
| 2.2.2 Inclusion/exclusion criteria . . . . .                            | 26        |
| 2.2.3 Data collection . . . . .   | 26        |
| 2.3 Results . . . . .   | 27        |
| 2.3.1 Methods for testing the validity of proxies . . . . .             | 27        |
| 2.4 Discussion . . . . .  | 33        |
| 2.4.1 Levels of agreement between the patient and proxy . . . . .       | 33        |
| 2.4.2 Factors associated with the patient and proxy . . . . .           | 35        |
| 2.4.3 Factors associated with assessment . . . . .                      | 39        |
| 2.4.4 Factors associated with the validity studies . . . . .            | 44        |
| 2.4.5 Summary and conclusions . . . . .                                 | 46        |
| 2.5 Research aims . . . . .   | 47        |
| 2.5.1 Research question one . . . . .                                   | 48        |
| 2.5.2 Research question two . . . . .                                   | 48        |
| 2.5.3 Research question three . . . . .                                 | 49        |
| 2.5.4 Research question four. . . . .                                   | 49        |
| 2.5.5 Summary . . . . .   | 49        |
| <b>CHAPTER 3: Methods and procedures</b>                                | <b>51</b> |
| 3.1 Introduction . . . . .  | 51        |
| 3.2 Cognitive interviews . . . . .                                      | 51        |
| 3.3 Verbal protocols . . . . .  | 52        |
| 3.3.1 Verbal protocol procedure . . . . .                               | 53        |
| 3.3.2 The validity of verbal protocols. . . . .                         | 54        |
| 3.3.3 Limitations of verbal protocols . . . . .                         | 55        |
| 3.4 Analysis of verbal protocols . . . . .                              | 56        |
| 3.4.1 Coding verbal protocols. . . . .                                  | 56        |
| 3.4.2 Coding scheme. . . . .  | 57        |
| 3.4.3 Reliability of the coding. . . . .                                | 57        |
| 3.5 Content analysis . . . . .  | 58        |

|   |           |
|---|-----------|
| 3.5.1 Quantitative analysis of content . . . . .                                | 58        |
| 3.5.2 Qualitative content analysis . . . . .                                    | 59        |
| 3.5.3 Ethics of cognitive interviews.. . . .                                    | 59        |
| 3.6 Design and procedures . . . . .   | 60        |
| 3.6.1 Ethics of interviewing bereaved relatives . . . . .                       | 60        |
| 3.6.2 Sample . . . . .  | 61        |
| 3.6.3 Measures and procedure . . . . .  | 61        |
| 3.6.4 Pilot study . . . . .   | 63        |
| 3.6.5 Follow up study . . . . .   | 63        |
| <b>CHAPTER 4: Frequency and severity descriptors</b>                            | <b>65</b> |
| 4.1 Introduction . . . . .  | 65        |
| 4.1.1 Pain . . . . .  | 66        |
| 4.1.2 Anxiety and depression . . . . .  | 67        |
| 4.2 Aims . . . . .  | 69        |
| 4.3 Methods. . . . .  | 69        |
| 4.4 Analysis and findings . . . . .   | 69        |
| 4.4.1 Descriptive statistics . . . . .  | 70        |
| 4.5 Numerical data analysis of the severity and frequency descriptors . . . . . | 72        |
| 4.6 Qualitative analysis of the severity and frequency descriptors. . . . .     | 73        |
| 4.6.1 Findings from the pain severity descriptors. . . . .                      | 73        |
| 4.6.2 Findings from the anxiety severity descriptors . . . . .                  | 75        |
| 4.6.3 Findings from the depression severity descriptors. . . . .                | 77        |
| 4.6.4 Findings from the frequency descriptors . . . . .                         | 79        |
| 4.7 Discussion . . . . .  | 80        |
| 4.7.1 Descriptions of the pain descriptors. . . . .                             | 81        |
| 4.7.2 Descriptions of the anxiety and depression descriptors. . . . .           | 81        |
| 4.7.3 Limitations. . . . .  | 83        |
| 4.8 Conclusion. . . . .   | 83        |
| <b>CHAPTER 5: Proxies' symptom judgments</b>                                    | <b>85</b> |
| 5.1 Introduction . . . . .  | 85        |
| 5.2 The cognitive model of question answering and proxies . . . . .             | 85        |
| 5.2.1 Response difficulties . . . . .   | 85        |
| 5.2.2 Memory for events. . . . .  | 86        |
| 5.2.3 Cues and judgments. . . . .   | 86        |
| 5.2.4 Frequency judgments. . . . .  | 88        |
| 5.2.5 Reference periods and event cues . . . . .                                | 89        |
| 5.3 Aims . . . . .  | 90        |
| 5.4 Methods. . . . .  | 90        |
| 5.5 Findings and discussion . . . . .   | 91        |
| 5.5.1 Proxies' responses to VOICES . . . . .                                    | 91        |
| 5.5.2 Analysis of the PSCI . . . . .  | 94        |
| 5.5.3 Reliability/validity . . . . .  | 97        |
| 5.5.4 Response difficulties . . . . .   | 98        |
| 5.5.5 Symptom frequency judgments . . . . .                                     | 100       |
| 5.5.6 Symptom cues and judgments . . . . .                                      | 102       |
| 5.5.7 Event cues . . . . .  | 106       |
| 5.5.8 Reference periods. . . . .  | 107       |
| 5.6 Conclusion. . . . .   | 108       |
| 5.6.1 Limitations. . . . .  | 110       |
| 5.6.2 Summary . . . . .   | 110       |

|   |            |
|---|------------|
| <b>CHAPTER 6: Proxies’ emotions and judgments</b>                     | <b>112</b> |
| 6.1 Introduction . . . . .  | 112        |
| 6.1.1 Grief . . . . .   | 112        |
| 6.1.2 Cognition and emotions . . . . .                                | 114        |
| 6.2 Aims . . . . .  | 115        |
| 6.3 Methods . . . . .   | 116        |
| 6.4 Results . . . . .   | 116        |
| 6.4.1 Descriptive statistics . . . . .                                | 116        |
| 6.4.2 Analysis procedure . . . . .                                    | 117        |
| 6.5 Discussion . . . . .  | 121        |
| 6.5.1 Limitations . . . . .   | 123        |
| 6.5.2 Summary . . . . .   | 123        |
| <b>CHAPTER 7: Proxies’ responses over time</b>                        | <b>125</b> |
| 7.1 Introduction . . . . .  | 125        |
| 7.1.1 Changes over time . . . . .                                     | 125        |
| 7.1.2 Timing of interviews . . . . .                                  | 126        |
| 7.1.3 Factors likely to affect proxies’ responses over time . . . . . | 127        |
| 7.2 Aims . . . . .  | 128        |
| 7.3 Methods . . . . .   | 129        |
| 7.4 Results . . . . .   | 129        |
| 7.4.1 Descriptive statistics . . . . .                                | 129        |
| 7.4.2 Proxy self-report measures . . . . .                            | 131        |
| 7.4.3 Analysis procedure . . . . .                                    | 132        |
| 7.4.4 Quantitative differences between time 1 and time 2 . . . . .    | 132        |
| 7.4.5 Proxies’ responses to VOICES . . . . .                          | 134        |
| 7.4.6 Qualitative analysis . . . . .                                  | 139        |
| 7.5 Discussion . . . . .  | 145        |
| <b>CHAPTER 8: Discussion</b>  | <b>148</b> |
| 8.1 Research aims and findings . . . . .                              | 148        |
| 8.1.1 Research question one . . . . .                                 | 148        |
| 8.1.2 Research question two . . . . .                                 | 149        |
| 8.1.3 Research question three . . . . .                               | 152        |
| 8.1.4 Research question four . . . . .                                | 153        |
| 8.1.5 Contributions to the field of research . . . . .                | 155        |
| 8.2 Future research . . . . .   | 157        |
| 8.3 Limitations . . . . .   | 159        |
| 8.3.1 Sample size . . . . .   | 159        |
| 8.3.2 Methods and design . . . . .                                    | 160        |
| 8.4 Interviewing bereaved relatives . . . . .                         | 162        |
| 8.5 Conclusion . . . . .  | 162        |
| <b>References</b>   | <b>164</b> |
| <b>List of Publications</b>   | <b>187</b> |

|                    |   |            |
|--------------------|---|------------|
| <b>Appendix A:</b> | <b>Factors likely to affect the validity of proxies’ retrospective reports .</b>  | <b>188</b> |
| <b>Appendix B:</b> | <b>Ethics protocol .....</b>  | <b>190</b> |
| <b>Appendix C:</b> | <b>Recruitment letter and participant consent for study one.....</b>  | <b>192</b> |
| <b>Appendix D:</b> | <b>Demographic and patient/proxy characteristics.....</b>   | <b>196</b> |
| <b>Appendix E:</b> | <b>Caregiver Strain Index .....</b>   | <b>199</b> |
| <b>Appendix F:</b> | <b>Proxy Semi-structured Cognitive Interview .....</b>  | <b>201</b> |
| <b>Appendix G:</b> | <b>Differences between questions in the proxy semi-structured cognitive interview and questions from the retrospective palliative care questionnaires .....</b> | <b>215</b> |
| <b>Appendix H:</b> | <b>Prompts used in the Proxy Semi-structured Cognitive Interview ...</b>  | <b>218</b> |
| <b>Appendix I:</b> | <b>Modifications to the Proxy Semi-structured Cognitive Interview following the pilot study .....</b>   | <b>221</b> |
| <b>Appendix J:</b> | <b>Recruitment letter and participant consent for the follow up study..</b>   | <b>223</b> |
| <b>Appendix K:</b> | <b>Shortened version of the Proxy Semi-structured Cognitive Interview used in the follow up study .....</b>   | <b>226</b> |
| <b>Appendix L:</b> | <b>Questions relating to symptom severity and frequency descriptors..</b>   | <b>236</b> |
| <b>Appendix M:</b> | <b>Differences between individuals who took part in the study and those who did not.....</b>  | <b>239</b> |
| <b>Appendix N:</b> | <b>Coding frequencies of the Proxy Semi-structured Cognitive Interview.....</b>   | <b>241</b> |
| <b>Appendix O:</b> | <b>Analysis of differences in VOICES ratings between those with high and low levels of self-report caregiver strain .....</b>                                   | <b>251</b> |
| <b>Appendix P:</b> | <b>Analysis of differences between proxies who were followed-up and those who were not. ....</b>  | <b>253</b> |

# List of Figures

|            |   |     |
|------------|---|-----|
| Figure 3.1 | Summary of research design. . . . .   | 64  |
| Figure 4.1 | Category frequencies for the pain intensity descriptors.. . . .   | 74  |
| Figure 4.2 | Category frequencies for the anxiety intensity descriptors. . . . .   | 76  |
| Figure 4.3 | Category frequencies for the depression intensity descriptors. . . . .  | 77  |
| Figure 5.1 | Proxies’ ratings of patients’ pain severity in the last week and month of life. . . . .                             | 91  |
| Figure 5.2 | Proxies’ ratings of patients’ pain frequency in the last week and month of life. . . . .                            | 92  |
| Figure 5.3 | Proxies’ ratings of patients’ anxiety severity in the last week and month of life. . . . .                          | 93  |
| Figure 5.4 | Proxies’ ratings of patients’ anxiety frequency in the last week and month of life. . . . .                         | 93  |
| Figure 5.5 | Proxies’ ratings of patients’ depression severity in the last week and month of life. . . . .                       | 94  |
| Figure 5.6 | Proxies’ ratings of patients’ depression frequency in the last week and month of life.. . . .                       | 94  |
| Figure 7.1 | Grouping of VOICES frequency and severity descriptors. . . . .  | 133 |
| Figure 7.2 | Proxies’ ratings of patients’ pain frequency in the last week and month of life at time 1 and time 2.. . . .        | 135 |
| Figure 7.3 | Proxies’ ratings of patients’ pain severity in the last week and month of life at time 1 and time 2. . . . .        | 135 |
| Figure 7.4 | Proxies’ ratings of patients’ anxiety frequency in the last week and month of life at time 1 and time 2.. . . .     | 136 |
| Figure 7.5 | Proxies’ ratings of patients’ anxiety severity in the last week and month of life at time 1 and time 2.. . . .      | 137 |
| Figure 7.6 | Proxies’ ratings of patients’ depression frequency in the last week and month of life at time 1 and time 2. . . . . | 138 |
| Figure 7.7 | Proxies’ ratings of patients’ depression severity in the last week and month of life at time 1 and time 2.. . . .   | 138 |
| Figure 7.8 | Categories and themes from qualitative analysis.. . . .   | 140 |

# List of Tables

|            |   |     |
|------------|---|-----|
| Table 2.1: | Prospective studies comparing patients and their significant other's ratings on aspects of the patients' experiences relevant to quality of care at the end of life. ....       | 28  |
| Table 2.2: | Prospective/retrospective studies comparing patients and their significant other's ratings the quality of care at the end of life. ....   | 33  |
| Table 4.1: | Patient demographics and characteristics. ....  | 70  |
| Table 4.2: | Proxy demographics and characteristics. ....  | 72  |
| Table 4.3: | Numerical ratings of severity adjectives for symptom severity in VOICES...  | 73  |
| Table 5.1: | Coding scheme for PSCI. ....  | 95  |
| Table 6.1: | Content analysis codes for proxies' emotional reactions and affect-based judgments.....   | 116 |
| Table 6.2: | The absolute frequencies of codes for CSI. ....   | 119 |
| Table 6.3: | The absolute frequencies of codes for SAI. ....   | 120 |
| Table 6.4: | Analysis of differences in VOICES ratings between those with high and low levels of self-reported anxiety.....  | 120 |
| Table 7.1: | Patient demographics and characteristics. ....  | 129 |
| Table 7.2: | Proxy demographics and characteristics. ....  | 130 |
| Table 7.3: | Analysis of differences between proxies who took part in the follow up (time 2) and those who did not in self-reported levels of carergiver strain, anxiety and depression..... | 132 |
| Table 7.4: | Kappa measures of agreement for proxies' ratings of patients' pain at time 1 and time 2. ....   | 134 |
| Table 7.5: | Kappa measures of agreement for proxies' ratings of patients' anxiety at time 1 and time 2. ....  | 136 |
| Table 7.6: | Kappa measures of agreement for proxies' ratings of patients' depression at time 1 and time 2. ....   | 137 |

# CHAPTER 1: Introduction

The provision of hospice and palliative care for patients at the end of life is growing. Recent estimates in the UK show that approximately 19% of patients with cancer die in a hospice and approximately 60% of cancer patients who die are cared for by palliative home care teams (The Hospice Information Service, 2000). Palliative care is a multidisciplinary specialty, whose aims are,

*“the active total care of patients whose disease is not responsive to curative treatment. Control of pain, of other symptoms, and of psychological, social and spiritual problems is paramount. The goal of palliative care is achievement of the best quality of life for patients and their families.”*

(World Health Organization technical report series 804, 1990, p.11)

As this definition implies, palliative care goes beyond purely symptom control to incorporate quality of life, physical, spiritual and emotional aspects. This is commonly referred to as a holistic approach to the care of the individual and their family. This requires the skills of a variety of different professionals, including doctors, nurses, social workers, chaplains, therapists and sometimes psychologists or psychiatrists.

With the prevailing ethos for evidence-based practice and economic constraints, the provision of palliative care, like other areas of health care, has had to respond to demands for efficiency and effectiveness. The implication of this has been a move toward greater evaluation of services and to clinical audit. Evaluation and audit within palliative care serves various essential functions beyond economic concerns. These include identifying patient and family needs, assessing palliative care outcomes such as quality of life, monitoring standards and identifying training needs (Higginson, 1993). Thus, the methodologies employed must be both reliable and valid if interventions and policies are to be implemented in response to findings.

This chapter presents a brief introduction to some of the methodological challenges associated with evaluation and audit of services for the dying and their families. Since this Ph.D. is concerned with the *retrospective* or *after death* approach, the emphasis will be on this method. In retrospective palliative care studies, other people (family, friends, and health professionals), commonly referred to as proxies, respond on behalf of the patient. Most of the criticisms levelled at this approach have centred on the use of these proxies and how accurately their responses reflect the patients' experiences (Higginson *et al.*, 1994; Hinton, 1996). In order to better understand the issues, this chapter examines the use of proxies in other areas of health. By expanding the review to include a wider range of literature, an insight into factors likely to affect the validity of proxies' reports may be identified. This will also provide a



basis upon which to critically evaluate the literature on the use of significant others as proxies for patients at the end of life: the focus of Chapter 2.

## **1.1 Methodological problems when evaluating care of the dying**

When evaluating services for patients at the end of life and their families, researchers are confronted with a number of methodological challenges. One such problem is the difficulty in ascertaining the views and experiences of terminally ill patients. Apart from ethical concerns about interviewing patients, studies relying on patients' accounts prior to death, the prospective approach, are potentially biased as they only represent a proportion of the population that is less impaired and therefore able to participate, and willing to take part. A decline in performance status is seen at the end of life in many cancer patients and is a good indicator of prognosis (Vigano *et al.*, 2000). It is at a point in time that the assessment of patients' quality of life are most needed (Aaronson, 1991). Prospective studies are also limited to diseases such as cancer that have an identifiable terminal phase, and where patients are likely to be known to relevant services (Morris *et al.*, 1986). Added to this, high attrition rates, small sample sizes and difficulties associated with predicting survival rates can also hamper attempts to conduct valid research (McWhinney *et al.*, 1994; Rinck *et al.*, 1997).

## **1.2 Retrospective approach**

An alternative approach, that overcomes many of the problems associated with prospective studies, is the retrospective or after death approach. Here, observations are gathered from proxies acting on behalf of the patient, after the patient's death. Usually this is the patient's next-of-kin or a health professional involved in the patient's care. This approach has been used in a number of influential studies on the care of the dying (Cartwright & Seale, 1990; Seale & Cartwright, 1994; Addington-Hall & McCarthy, 1995a). The largest survey of its kind in the United Kingdom is the Regional Study of the Care of the Dying (RSCD) (Addington-Hall & McCarthy, 1995a). Being retrospective, the authors were able to identify those individuals who had died within a specific time period; in this case the last quarter of 1990. From these, representative samples of 270 deaths from each of twenty different health districts in England were extracted. By selecting patients based solely on time and place of death, there is no inherent bias towards any particular cause of death. Thus, deaths from different causes could be represented (Addington-Hall *et al.*, 1995; McCarthy *et al.*, 1996; McCarthy *et al.*, 1997). The results of the RSCD show that the current health care and social services in England are in need of significant improvement before they can fully meet the needs of terminally ill patients and their families. Similar methods have been used in other countries, for example in the SUPPORT study where bereaved relatives provided information

on patients' last three days of life (Lynn *et al.*, 1997). These studies have provided important insights into death and dying in late 20th century.

However, collecting information indirectly from proxies at some later point in time introduces potential problems regarding the validity of reports (Higginson *et al.*, 1994; Hinton, 1996). Validity, in this context, refers to the degree to which the proxy's account accurately reflect the experiences of the individual they are representing. At present, little is known about the accuracy of proxies' responses, despite the fact that inaccuracy may compromise validity and lead to erroneous conclusions. Given that proxies are a significant source of information and are often the only available source for a large proportion of dying patients, it is vital to understand how and why their accounts might differ from those of the patient. The validity of proxies in other areas of health has been evaluated. Therefore, before critically evaluating the literature it is important to examine the use of proxies in other areas of health. The discussion begins with the validity of significant other proxies in areas of health and then moves on to examine health professional as proxies.

### **1.3 Proxies in other areas of health**

Similarly to palliative care, proxies report on the experiences of those who are not able to report for themselves such as children, those with learning difficulties, cognitive and emotional impairments. Proxies can also be used to corroborate information provided by self-report when there is some reason to doubt the accuracy of self-report, for instance when reporting on habits such as smoking, drug and alcohol consumption. Another area where proxies are widely used is in epidemiology research. In this instance an alternative source is used when the individual is not available or unable to give a self-report. More often than not, this is another member of the same household. Research that has examined the validity of proxies in these areas will be discussed, with the exception of studies that concern patient-proxy agreement involving children, adolescents and individuals with learning difficulties. These groups have their own specific needs as indicated in recent reviews examining the use of proxies with these groups (Stancliffe, 2000; Eiser & Morse, 2001a; 2001b).

### **1.4 Significant others as proxies**

#### **1.4.1 Epidemiology**

Epidemiology field studies often rely on other family members to provide information on morbidity, demographics and behaviours such as physical activity, smoking, diet, and alcohol consumption in exposure-disease research. Proxies are used because interviewing the individuals themselves may not be feasible, for example in individuals with conditions that are rapidly fatal, cognitively or communicatively compromising and with some psychiatric

conditions (Nelson *et al.*, 1990). In studies of morbidity and mortality information is sometimes sought from proxies after the patient's death (Rogot & Reid, 1975; Pickle *et al.*, 1983; Lerchen & Samet, 1986). Proxies also provide researchers with more representative samples, as Kolonel *et al.* (1977) points out, samples would consist of only those individuals at home who are available for interview, such as housewives and the elderly. Moreover, it is less time consuming and more economical to interview proxies.

Given the importance of the information in identifying diseases and the links between disease and exposure to hazards, there have been numerous studies examining the reliability of proxy data (see Nelson *et al.*, 1990; for review). Comparisons between proxies and absent household members and/or medical judgments suggests that proxies can provide reliable data with regard to certain types of information. Several studies have shown proxies' reports are reliable for factual information such as age, height, occupational history (Rogot & Reid, 1975; Hatch *et al.*, 1991; Colt *et al.*, 2001), smoking status (Rogot & Reid, 1975; Kolonel *et al.*, 1977; Herrmann, 1985; Lerchen & Samet, 1986; Emont *et al.*, 1991; Hatch *et al.*, 1991; Halabi *et al.*, 1992), dietary intake (Moore *et al.*, 1970; Kolonel *et al.*, 1977; Lerchen & Samet, 1986), health care utilisation (Mosley & Wolinsky, 1986) and certain health conditions (Halabi *et al.*, 1992). However, the findings have been variable with differing levels of congruency and poorer levels of concordance in reporting some aspects. Among these are alcohol consumption (Hatch *et al.*, 1991), occupational history (Lerchen & Samet, 1986), some health conditions (Halabi *et al.*, 1992), pregnancy outcomes in reproductive studies (Fickree *et al.*, 1993), dietary history (Herrmann, 1985) and number of cigarettes smoked (Lerchen & Samet, 1986). This has led some investigators to question the use of proxies.

Several factors appear to contribute to the accuracy of proxy reporting in epidemiology studies. Evidence suggests that the type of information sought from proxies is important. Halabi *et al.* (1992) found that proxies were reliable for providing information on conditions that were salient, such as heart disease. Whereas, agreement between patients, proxies and medical examination were lowest for back pain. Correspondingly, Claridge and Massagli (1989), in a study of female proxies, found that agreement on the presence of twenty-two physical and psychological complaints was better as the persistence and thus the saliency of the problem increased. Proxies in this study were found to underestimate self-reported complaints, which the authors suggest was due to the lack of gravity given the type of complaints (e.g. cough, sinus trouble, rash). Claridge and Massagli's (1989) study was also significant because it showed that there was a poor level of agreement for the less observable and more subjective psychological complaints of nervousness and depression.

The type of proxy has important repercussions regarding how the information sought by the researcher is affected. In a study of exposure-mortality, Pickle *et al.* (1983) found that

siblings were better able to respond to questions about the deceased individual's immediate family and events occurring early in life, while spouses and offspring were best able to describe events that occurred during adult life. Alternative proxies, such as more distant relatives and friends were the poorest at responding to questions. Other studies have shown that the reliability of proxies is related to the person who is acting as the proxy. Emont *et al.* (1991) used proxies to corroborate self-reported smoking cessation and found that discrepancies in accounts were influenced by the opportunities the proxy had to observe the individual in different social contexts. These findings highlight the importance of the proxy having access to the relevant information. Naturally, proxies who reside together are usually the most reliable source and are likely to provide the most complete information (Emont *et al.* 1991; Anonymous, 2000).

The ability of the proxy to provide information also appears to be influenced by the level of detail required. Pickle *et al.* (1983) found that the completeness of the information was affected by the amount of detail requested. For instance, proxies were more accurate at stating whether the patient smoked but were less accurate at giving the number of cigarettes smoked (Lerchen & Samet, 1986). With regard to characteristics of the patient and proxy there has been little consistency in the findings with some studies identifying factors (Pickle *et al.*, 1983; Emont *et al.*, 1991; Hatch *et al.*, 1991) and others finding no association (Kolonel *et al.*, 1977; Herrmann, 1985).

### **1.4.2 Quality of life**

The use of proxies allows patients who would otherwise be excluded to take part in studies and thus improves the validity and generalisability of the research. However, more importantly, accurate assessment of patients' quality of life by significant others is vital in monitoring their condition and its management. Since many relatives are also carers for the patient, they are often called upon to be proxies. For this reason, the validity of significant other proxies for patients with acute and chronic conditions, cognitive/intellectual and communicative impairments and with older individuals is important and has been extensively investigated. In the following overview of the subject, for ease of clarity the subsections are divided but in reality there is often overlap between the areas. Consequently there is some overlap in the studies.

#### ***Acute & chronic health conditions***

Proxies have been employed to report on behalf of a number of patients with acute and chronic health conditions. These include trauma, surgical and medical patients in hospitals and the community receiving treatment (Sneeuw *et al.*, 1992; Rogers *et al.*, 1997; Capuzzo *et*

*al.*, 2000). To date, the investigations conducted have been with chronic health conditions such as cancer, heart and lung diseases have been with older individuals (e.g. Magaziner *et al.*, 1988; Epstein *et al.*, 1989). Few studies have specifically evaluated proxies in younger adult populations or in the context of critical and acute health care settings. Within the context of intensive care studies the evidence has been contradictory. Capuzzo *et al.* (2000) found that although accuracy of emotional aspects of quality of life were lower, agreement between patients' and their relatives' ratings of quality of life prior to admission to the unit were excellent. The authors concluded from these findings that proxies can be acceptable substitutes for patients. Contrary to this, Rogers *et al.* (1997) comparing patients' and relatives' ratings of quality of life at discharge and six months after discharge from a critical care unit, found that overall agreement was weak, with some improvement at six months follow-up. Again, reliability was poorest for mental health. Differences in the timing of the assessment and the patient's health could have contributed to the inconsistencies observed between these studies.

### ***Older individuals***

In older individuals cognitive, physical and sensory deficits may prevent them providing self-reports or information on some aspects of their functioning. Limitations that are not uncommon in this group (Magaziner *et al.*, 1988; Evans *et al.*, 1990). Studies of older individuals' functioning have in general been consistent in finding that agreement between the patient and proxy is good for more observable aspects of functioning. These include activities of daily living (e.g. bathing, dressing), physical activity, and conditions with more observable manifestations, or as Magaziner *et al.*, (1996) points out conditions that are likely to be discussed or are more serious such as diabetes and cancer (Rubenstein *et al.*, 1984; Epstein *et al.*, 1989; Magaziner *et al.*, 1988; Weinberger, *et al.*, 1992; Segal & Schall, 1994; Magaziner *et al.*, 1996; 1997; Long *et al.*, 1998; Okamoto, 2000). While agreement is lower for less observable and more subjective aspects, such as emotional status (Rubenstein *et al.*, 1984; Kiyak *et al.*, 1992; Weinberger, *et al.*, 1992; Segal & Schall, 1994; Magaziner *et al.*, 1997; Okamoto, 2000), less overt conditions like arthritis and hypertension (Farrow & Samnet, 1990; Magaziner *et al.*, 1992; 1996) and symptoms like pain (O'Brien & Francis, 1988; Werner *et al.*, 1998). Only moderate correlations between elderly individuals' and their relatives' ratings of pain have been reported (Werner *et al.*, 1998). Agreement on measures of quality of life were also low to moderate (Slevin, 1988; Sprangers & Aaronson, 1992).

Bias toward reporting greater disability (Rubenstein *et al.*, 1984; Epstein *et al.*, 1989; Rothman, *et al.*, 1991; Magaziner *et al.*, 1992; Weinberger, *et al.*, 1992; Magaziner *et al.*, 1997), pain (Werner *et al.*, 1998), poorer quality of life (Slevin *et al.*, 1988), pain (Ell *et al.*, 1988), cognitive function (Bassett *et al.* 1990) and psychological health (Epstein *et al.*, 1989;

Magaziner *et al.*, 1996; 1997), have been shown. One explanation for the tendency of proxies to over- and underestimate, concerns the role of caregiving. There is evidence that family members' psychological distress and/or perceived caregiver burden are related to the accuracy of assessment of patients' impairments, with greater burden associated with patients being rated as more impaired (Magaziner *et al.* 1988; Rothman *et al.*, 1991; Long *et al.*, 1998). Other factors such as the frequency of contact and the patient's health condition have also been shown to affect the level of concordance (Epstein *et al.*, 1989). Magaziner *et al.* (1988) found that agreement was better when the patient was not cognitively impaired or depressed. Okamoto (2000) identified spouses as being more accurate than other family members, such as children, in a study of elderly individuals living in the community.

### ***Cognitive/intellectual and communicative impairments***

Neurological conditions, such as Alzheimer's disease (AD), epilepsy, cerebro-vascular accident (CVA) and traumatic brain injury (TBI) can impair cognitive functioning and communication. The vast majority of the research has been conducted on patients with AD and related conditions where family members are used as proxies. Few have specifically examined the validity of proxies for patients with epilepsy (Hays *et al.*, 1995), CVA (Segal & Schall, 1994; Sneeuw *et al.*, 1997c; Duncan *et al.*, 2002) and TBI (Cusick, *et al.*, 2000). Generally agreement is better for those individuals who have higher cognitive functioning (Magaziner *et al.*, 1988; Kiyak *et al.*, 1994; Werner *et al.*, 1998; Loewenstein *et al.*, 2001; Novella, *et al.*, 2001), or are less severely affected in CVA patients (Sneeuw *et al.*, 1997c; Duncan *et al.*, 2002). Kiyak *et al.* (1994) comparing family members' reports with those from older demented patients' and non-demented patients' reports, found that those with dementia were rated as more impaired. As a result levels of agreement vary depending, to some extent, on the population being studied. For example poor to moderate agreement has been found for self-report and performance based measures of functioning and disability in AD patients (Loewenstein *et al.*, 1991; Novella, *et al.*, 2001). While, other studies have found high levels of agreement for functioning that is more concrete and observable, like physical functioning, between proxies and patients with TBI (Cusick *et al.*, 2000), CVA (Segal & Schall, 1994; Snueew *et al.*, 1997c; Duncan *et al.*, 2002) and AD (Zanetti *et al.*, 1999). Once again aspects of a subjective nature such as psychosocial functioning, thinking and memory were less congruent with patients' ratings (Segal & Schall, 1994; Snueew *et al.*, 1997c; Duncan *et al.*, 2002).

A consistent finding is that assessment of subjective aspects of patients' functioning and quality of life have been less concordant. Proxies' estimates of patients' cognitive functioning has been shown to be poor (Cusick *et al.*, 2000) and overestimated in some studies (Bassett *et al.*, 1990; Hays *et al.*, 1995). Proxies' assessments of patients' quality of life and

subjective aspects such as pain, emotions, social and role limitations also diverge from the patients (Segal & Schall, 1994). A number of biases have been identified. These include underestimates of patients' quality of life in AD patients (Novella, *et al.*, 2001) and overestimates of levels of depression in patients with dementia (MacKenzie *et al.*, 1989; Novella, *et al.*, 2001). Proxies also report lower levels of functioning than the patients themselves (Kiyak *et al.*, 1994; Sneeuw *et al.*, 1997c). However, proxies have been shown to overestimate impaired dementia patients' objective functional performance (i.e. tell the time, identify currency) (Loewenstein *et al.*, 2001).

One explanation for the inaccuracies observed relates to the patient's neurological condition and the symptom being assessed. Some symptoms may be problematic because they are difficult to differentiate in patients with cognitive impairment. For example, overlap between depression and dementia symptoms such as difficulty concentrating, and loss of interest are thought to have contributed to family members' overestimations of dementia patients' depression ratings (MacKenzie *et al.*, 1989). Teri and Truax (1994) examined whether family caregivers own depression was reflected in their ratings of patient depression. Although the proxies' depression was found to be moderately linked to their ratings of patients' depression, there was no substantial association. Thus the authors' concluded that caregivers were not biased on account of their own mood. However, it is important to note that the levels of proxy depression were mild in this study. Despite this finding, characteristics of the proxy have been shown to contribute to discrepancies in their ratings. In a study of very mild and mild dementia, the family members' perceived level of caregiver burden was associated with differences in patients' and proxies' ratings of activities of daily living and physical functioning (Zanetti *et al.*, 1999). Consistent with the finding in older individuals (Okamoto, 2000), spouses were found to be more accurate than other family members (Novella, *et al.*, 2001).

In assessing quality of life in patients with acute and chronic conditions, older individuals and patients with cognitive/intellectual and communicative impairments, there has been no other consistent relationships demonstrated between demographics associated with either the patient or proxy, even though a number of studies have found factors. For example, patient and proxy age, (Rothman *et al.*, 1991; Magaziner *et al.*, 1996), educational attainment (McCusker & Stoddard, 1984; Hays *et al.*, 1995; Magaziner *et al.*, 1996) and gender (Magaziner *et al.*, 1996). Variations in the populations, context, methods and measures could account for this. Consistent with the epidemiology studies, the congruency between the patient and proxy in quality of life assessment varies according to the measure. Complex or ambiguous questions can lead to different interpretations of the same question (Magaziner *et al.*, 1988; 1996). The number of response options is also important, as scoring is better for

dichotomous responses than for multiple response options (Magaziner *et al.*, 1997). In addition to this, the reliability of the scale can influence the level of agreement (Snueew *et al.*, 1997).

### **1.4.3 Health care decisions**

When patients are incapable of making judgements about their care and treatment, proxies are sometimes used to make decisions on their behalf. There is increasing interest in this area with advance directives (Aikman *et al.*, 1999; Volicer *et al.*, 2002). More often than not, proxies' and patients' decisions regarding treatments are compared in hypothetical clinical situations. For example, Seckler *et al.* (1991) investigated the accuracy of proxies' decisions compared to chronically ill elderly patients' wishes in hypothetical situations regarding cardiopulmonary resuscitation (CPR). One scenario was in present health and another scenario under circumstances of progressive dementia. Agreement was poor with no significant direction to discrepant responses (to provide or withhold treatment). Another study found that proxies' overestimated patients' preferences for CPR (Uhlmann *et al.*, 1988). The findings have raised questions about the use of proxies given that family members have not adequately predicted patients' decisions regarding treatment decisions in a number of studies (Uhlmann *et al.*, 1988; Diamond *et al.*, 1989; Ouslander *et al.*, 1989; Seckler *et al.*, 1991; Hare *et al.*, 1992). Although there are exceptions with agreement on treatment (Libbus & Russell, 1995; Potkins, *et al.*, 2000) and care setting preferences (Mattimore *et al.*, 1997).

To further investigate the accuracy of substituted judgments, Sulmasy *et al.* (1998) examined factors associated with levels of accuracy in patients with a terminal diagnosis and medical patients. Accuracy was associated with the treatment and the scenario. In permanent coma scenario accuracy was better than when there was coma with a chance of recovery, and severe dementia. In addition to this, multiple factors were related to congruency. Agreement was higher when there was explicit discussion between the patient and proxy about their preferences, in better educated patients and proxies, and when the patient had private health insurance. Conversely, a diagnosis of heart failure, the proxy's experience of life-sustaining treatment, the proxy's involvement with religious services and the patient's belief that they would live longer than ten years, was negatively related to agreement. An understanding of factors that influence agreement will be useful for identifying conditions under which accuracy is likely so that ways to improve it can be instigated (Sulmasy *et al.*, 1998).

## **1.5 Health professionals as proxies**

Health professionals as part of their role, routinely use their judgment to make decisions about the patient's condition, treatment and care. As Zanetti *et al.* (1999) points out, accurate



assessment is necessary for monitoring the patient's condition and for planning and implementing appropriate care. For the reasons outlined earlier, (Section 1.3) patients are not always capable or willing to express themselves. This section will discuss the validity of health professionals as proxies when making health care decisions and assessing quality of life.

### **1.5.1 Health care decisions**

The majority of studies that have assessed the accuracy of health professionals' decisions as substitutes for the patient, have compared doctors and patients using scenarios involving threatening situations and treatments. Evidence from these studies is consistent in finding that doctors often do not reflect patients' wishes regarding treatment (Uhlmann *et al.* 1988; Zweibel & Cassel, 1989; Seckler *et al.*, 1991). In fact, Seckler *et al.* (1991) found that doctors did no better than chance. Others studies have found more positive results. Examining choices for nursing home residence, Mattimore *et al.* (1997) found that compared to significant others as proxies, doctors' ratings were more divergent. However, overall agreement was relatively high for both types of proxies. Another study found that preferences for CPR were congruent but varied according to the decision. Agreement was better when preferences which were normative, as opposed to unusual or contentious preferences (Layde *et al.*, 1995). For example, agreement was higher for those patients choosing CPR than for patients not choosing CPR (Layde *et al.*, 1995). Thus far, the evidence implies that doctors can not be relied upon to approximate patients' wishes. One explanation for the disparity may be the difficulties in predicting patients' health values. This is a problem identified by Tsevat *et al.* (1995; 1998), where seriously ill patients' rated their health status higher than their doctors. Added to this, the patient's health status did not predict their health values. If doctors' decisions are based on health status they are likely to be inaccurate. Consequently, a better understanding of patients' health values may be a measure to determining accurate preferences (Tsevat *et al.*, 1995).

### **1.5.2 Quality of life**

The vast majority of the research using health professionals as proxies has involved assessing quality of life in patients with chronic conditions such as cancer, and for patients at the end of life. The evidence suggests that doctors' and nurses' ratings diverge from those of the patients. Doctors' ratings of quality of life have been found to be poor to moderately congruent with the patients' ratings (Pearlman & Uhlmann, 1988; Slevin *et al.*, 1988; Sprangers & Aaronson, 1992; Kivinen *et al.*, 1998), although higher levels have been reported (Sprangers & Aaronson, 1992). In line with the research on significant others as proxies, agreement is higher for physical symptoms and functional items rather than psychosocial symptoms

(Brunelli *et al.*, 1998). Symptoms that are more subjective in nature, such as pain, anxiety, depression, drowsiness and nausea differ significantly between doctors and patients (Grossman *et al.*, 1991; Sprangers & Aaronson, 1992; Cleeland *et al.*, 1994; Grassi *et al.*, 1996; Brunelli *et al.*, 1998; Nekolaichuk *et al.*, 1999a). There is definitely a tendency for doctors to underestimate patients' quality of life (Pearlman & Uhlmann, 1988; Sprangers & Aaronson, 1992; Grassi *et al.*, 1996), problems (Brunelli *et al.*, 1998), symptoms (Derogatis *et al.*, 1976; Teri & Wagner, 1991; Nekolaichuk *et al.*, 1999a) and pain (Peteet *et al.*, 1986; Au *et al.*, 1994; Nekolaichuk *et al.*, 1999a). Although overestimates of quality of life (Jachuck *et al.*, 1982), anxiety (Derogatis *et al.*, 1976; Sensky *et al.*, 1989; Higginson & McCarthy, 1993) and problems (Higginson & McCarthy, 1993) have been found. In one study doctors' ratings of patients' quality of life were seen to improve over time (Grassi *et al.*, 1996). In contrast symptom assessment by doctors did not, with symptom assessments significantly lower than the patients on both assessment times (Nekolaichuk *et al.*, 1999a).

The findings have important implications for managing patients, as Cleeland *et al.*, (1994) identifies, discrepancies between patients' and doctors' estimations of pain are predictive of poor pain management. Indeed, the level of congruency has been found to be lower in patients' experiencing significant pain (Grossman *et al.*, 1991). Several explanations have been proposed for the observed differences. Within the context of acute medical settings, Jachuck *et al.* (1982), explains the lack of agreement as the dearth of objective clinical manifestations to measure subjective feelings, shorter consultations, inadequate communication and poor patient insight. Grossman *et al.* (1991) who reported that doctors', medical oncology fellows' and nurses' perceptions of pain intensity were not the same as the patients, implies that concern with issues, such as addiction and side effects, explains the disparities. However, many of the studies have taken place in palliative care settings where there is more contact between the patient and doctor. Moreover, the doctors are often oncologists who specialise in the area. What the studies do reveal is the complexity of quality of life and the inherent difficulties doctors contend with when assessing particular aspects.

Nurses' assessments as proxies have centred on patients' suffering with regard to pain and psychological distress. Estimates of pain show that nurses underestimate pain and problems associated with pain control (Grossman *et al.*, 1991; Higginson & McCarthy, 1993; Au *et al.*, 1994), and identify more problems than the patients themselves (Higginson & McCarthy, 1993). Again agreement is better for physical symptoms as opposed to those of a psychological nature (Holmes & Eburn, 1989; Peruselli *et al.*, 1992; Nekolaichuk *et al.*, 1999b). Accuracy of nurses' assessments of patients' symptoms, like doctors' assessments, have not been shown to improve over time (Nekolaichuk *et al.*, 1999a).

Measures of affective states (i.e. depression, anxiety) suggest that nurses, in general, overestimate the patients' suffering (Mason & Muhlenkamp, 1976; 1981; Husted & Johnson, 1985; Muhlenkamp & Joyner, 1986; Biley, 1989; Farrell, 1991; Higginson & McCarthy, 1993). Though underestimates of depression and anxiety have been found between nurses and patients in a palliative care setting (Nekolaichuk *et al.*, 1999b). A number of variables have been shown to influence nurses' assessments of patients' suffering. Davitz and Davitz (1981) in a series of investigations found that nurses' stereotype patients based on age, socio-economic status, gender, ethnic background and illness. For example, illnesses that were life threatening or where there was the possibility of long-term severe disability were considered psychologically distressing. A number of studies concur with Davitz and Davitz (1981) findings. Mason and Muhlenkamp (1976) in a study of nurse proxies for acutely ill patients who had undergone surgical amputation, found that nurses' rated patients as more anxious, depressed and hostile than the patients themselves. Attribution of anxiety by nurses has also been shown to be positively correlated with patients' social status, occupation and education in hospitalised arthritis patients (Mason & Joyner, 1985). While estimates of hostility ascribed to patients at the end of life were negatively correlated with patients' life expectancy (Husted and Johnson, 1985). There is also evidence in doctors' assessments that they stereotype based on patients' age and sex (Cleeland *et al.*, 1994; Kivinen *et al.*, 1998).

Jennings and Muhlenkamp (1981) suggest that in line with Wright's (1960) conceptualisation of *mourning*, nurses may view the patient as less fortunate and expect them to suffer in order to safeguard their own values of health and well-being. This may be useful in understanding the bias observed with mood states but does not explain nurses' underestimations of patients' pain. Furthermore, the bias to overestimate mood states may in some instances be attributable to the patient's own assessment. For example, denial by the patient has been shown to influence concordance (Jennings & Muhlenkamp, 1981; Husted and Johnson, 1985). Although the findings suggest that nurses' assessments of patients' suffering may be unreliable as a result of perceptual bias and stereotyping, others have failed to identify a relationship between the level of disagreement and characteristics of the patient or proxy. More work is therefore needed to examine when and where stereotyping is likely to occur and how to reduce its effects.

To date, only a few studies have compared agreement between significant others and health professionals with those of the patient. Higginson and McCarthy (1993) found that palliative care teams' ratings of patients' problems were closer to the patients than family members. Others have found that both spouses and nurses are more in line with the patient than other health professionals or family members (Novella, *et al.*, 2001). Nekolaichuk *et al.* (1999a), comparing health professionals' ratings with those of advanced cancer patients,

found that nurses' average ratings of nine symptoms, using the Edmonton Symptom Assessment System, were closer to patients' ratings than to the doctors' average ratings. Similar findings have been reported by Wilson *et al.* (2000) with proxies for advanced cancer patients. As nurses and significant others are likely to spend more time with the patient, they may have more opportunities to observe and assess the patient.

## 1.6 Discussion

Despite the diversity of the studies and areas evaluating the validity of proxies, there are consistent findings that emerge. Firstly, proxies need to have access to the information, either through interacting or observing the patient in order to make judgments. In the context of their home this is often a significant other who lives with the patient, while in hospital, nursing home or hospice it is likely to be nurses as they have more exposure to the patient. A caveat to this is the perceived burden of caring, which can bias proxies' perceptions of patients' quality of life; leading them to overestimate aspects of patients' functioning. Secondly, congruence is better for objective more observable phenomenon, such as physical functioning than subjective aspects like mood states and pain. The reliance on proxies for these less overt aspects may lead to mismanagement of some symptoms and spurious conclusions in research investigations. For example, the prevalence and severity of pain may be overestimated by significant others, while health professionals may underestimate it. This may have a bearing on conclusions regarding interventions and audit of services. Thirdly, the amount and type of detail requested has been identified as a factor that can affect levels of agreement. In several of the areas discussed information that is ambiguous or complex, or that requires the proxy to make specific judgments, such as frequency, are less congruent than those which are explicit and request less extensive information (i.e. absence or presence).

Within the different areas discussed there are also consistencies. In general, patients' levels of cognitive functioning influences agreement between significant others and patients. Other considerations include the patients' illness or condition, and the timing of the assessment. The literature investigating health professionals as proxies uncovers some degree of stereotyping when assessing patients. This could account for the differences observed, since assessments may be based on characteristics of the patient rather than observations. Stereotyping may be a product of being socialised within the context of the health care environment, where beliefs regarding patients' suffering may be shared. This would explain why agreement was related to nurses' experience (Jennings & Muhlenkamp, 1981; Husted & Johnson, 1985) and the stereotyping of older patients by doctors (Cleeland *et al.*, 1994; Kivinen *et al.*, 1998). Few decisive conclusions can be made regarding proxies when making health care decisions due to the dearth of literature. Only that significant others are more in line with the patient's

wishes than their doctors. Despite this the evidence does cast doubt on the reliability of proxies in this area.

In conclusion, the findings raise questions about the use of proxies, which if relevant to palliative care also raises important questions concerning the retrospective approach. Notwithstanding, the proxy may reliably report on less subjective aspects of the patients' experiences, which are also important elements in the evaluation of patient care. Whether the research discussed translates to significant others as proxies for patients at the end of life remains to be seen. This is the focus of the next chapter.

# CHAPTER 2: Proxy systematic review

## 2.1 Introduction

This chapter provides a comprehensive and critical review of studies that assess the validity of proxies acting for patients at the end of life. The emphasis is on significant others as proxies for the patient. These include family members, friends or neighbours, but exclude health professionals. This is because the focus of retrospective surveys reported to date, rely on significant others acting as proxies rather than health professionals<sup>1</sup>. Relevant literature from palliative and health care, survey methodology and psychology is drawn upon to examine factors likely to affect the validity of proxies' reports. From this, a broader insight into factors that can affect the validity of their reports may be revealed. In an attempt to explain the findings theories from social and cognitive psychology will be drawn upon. From this, a set of hypotheses will be posed as a means of investigating the validity of the retrospective approach further.

## 2.2 Method

### 2.2.1 Literature search strategy

A systematic review of the literature, using a recognised search strategy (Centre for Reviews and Dissemination, Report 4, 1996) was performed.

#### ***Databases***

British Nursing Index (1998-2002), CINAL (1982-2002), BIDS, MEDLINE (1980-2002), RCN Journals database (1985-2002), PSYCHINFO (1984-2002), EMBASE (1980-2002). The BIDS (Bath Information and Data Services) EMBASE, SOCIAL SCISEARCH (Social Sciences Citation Index) and IBSS (International Bibliography of the Social Sciences) databases were searched for the period 1993 to 2002 using the same search terms.

#### ***Search terms***

The search strategy used for these databases followed that suggested by CRD Report Number 4 (1996; Appendix A1) using the following words either singly or in combination:

*proxy#*, *surrogate#*, *advocate#*, *respondent#*, *next-of-kin#*, *carers#*, *health professionals#*, *spouse#* used singularly and exploded to include all subheadings

*palliative care* or *terminal* or *end-of-life* or *care of dying* or *hospice* or *bereavement* combined with each of the mesh terms (1.)

---

1. Proxy will be used to refer to significant other proxies unless otherwise stated.

*satisfaction, symptoms, emotions, pain, quality of life, quality of healthcare, attitude to death, health, depression, anxiety, grief, care* combined with each of the mesh terms (1.) and the results of search 2.

*evaluation or assessment or audit or survey or interview or outcome* combined with each of the mesh terms (1.) and the results of search 2 and search 3.

*retrospective, prospective, validity, reliability* combined with each of the mesh terms (1) and the results of search 2, 3 and 4.

Restrictions on the searches were that they were written in English, and were published between the dates specified above for each database. Subheadings not relevant to the area of interest were excluded in searches 2, 3, 4 and 5.

### ***Hand searches***

The following journals were searched by hand from the first date of issue 1990 to May 2002: British Medical Journal, European Journal of Cancer, Journal of Pain and Symptom Management, Journal of Palliative Care, Palliative Medicine, British Journal of Cancer, Social Science and Medicine

### ***The Internet***

Cancer sites on the Internet were also investigated for general and more up to date information on the most recent cancer publications. There are two main sites of use. The first is CancerWEB ([www.graylab.ac.uk](http://www.graylab.ac.uk) copyright 2002) which is put together by the National Cancer Institute. This site provided useful overview information but no additional references to those achieved from the other searches. The second site is OncoLink ([www.oncolink.upenn.edu](http://www.oncolink.upenn.edu) copyright 1994-2002, The trustees of the University of Pennsylvania).

## **2.2.2 Inclusion/exclusion criteria**

For the purpose of the review only studies that investigate outcomes relevant to care at the end of life, where at least some of the subjects have an advanced disease or a terminal illness, are included. Excluded from the review are studies that have compared patients and health professionals' reports at the end of life as these were discussed in Section 1.4. In addition to this, all those studies that concern patient-proxy agreement involving children and adolescents this is because this group have their own specific needs at the end of life.

## **2.2.3 Data collection**

From each study relevant to this review the following information was gathered: authors, year, country, study design, patient and proxy populations included, outcomes and outcome

measures, main findings in terms of levels of agreement and factors affecting the degree of agreement or bias.

## 2.3 Results

### 2.3.1 Methods for testing the validity of proxies

A review of the literature revealed that thirty six papers met the inclusion criteria. From the qualifying literature, a limited number of methods for testing the validity of proxy responses were identified. One approach is to compare patients' and proxies' views. In these studies, validity equates to the proxy's agreement with the patient's own responses. The majority of validity studies have used this approach to prospectively compare patients and significant others' responses to a number of outcomes including symptoms, quality of life, family well being and evaluations of care. Table 2.1 contains a summary of studies and their main findings. Studies are divided into cross-sectional designs that compare reports at a point in time or repeated measures designs, which make comparisons over time. Repeated measures designs have the advantage that trends in the levels of patient and proxy agreement can be interpreted in light of changes in the patient's condition. From this, extrapolations can be made to patient's who are more impaired (Sneeuw *et al.*, 1997a; Sneeuw *et al.*, 1997b). Furthermore, the validity of proxies reports can be examined in terms of responsiveness to changes in the patients condition over time, relative to the patients own reports impaired (Sneeuw *et al.*, 1997b; Sneeuw *et al.*, 1998).

As Table 2.1 shows, agreement is investigated at both the individual and group level. Individual level agreement exposes incongruency between patient-proxy dyads, which can be masked at the group level. This is the proportion of exact one-to-one agreement or approximate agreement. From this, factors associated with discrepancies can be distinguished. At the group level, systematic reporting biases (underestimation and overestimation) between patient and proxy agreement can be detected along with the magnitude of the bias. Besides this, groups can be divided based on various factors likely to influence patient and proxy agreement, such as the level of functioning (Sneeuw *et al.*, 1997b). Alternatively, groups can be divided based on the degree of congruency. In a study by Kurtz *et al.*, (1996) they discerned differences between patient and proxy characteristics among congruent and non-congruent patient-proxy dyads.

A less commonly used approach is to examine the patient and proxies ratings independently. This does not assume that the patient is the *gold standard* upon which the significance of proxies' reports can be assessed. This is an important consideration, as Sneeuw *et al.* (1998) acknowledge, patients' reports are themselves prone to bias. An observation that is supported in the literature (Pennebaker, 1983; 1984). If patients reports are biased it is



unlikely that patients and proxies reports will be similar. One method, by which this can be achieved, is to separate out the proportions of variance attributable to a score. If variables other than those expected account for a large proportion of the variance, then this suggests that the responses are not valid (Sneeuw *et al.*, 1998). Using this method, Sneeuw *et al.* (1998) found that both patient and proxy ratings of patients' quality of life were explained largely by patients' health status supporting the validity of their responses. The advantage of this approach is that the identity of factors that contributed to patient and proxy ratings can be extracted. For example, Elliott *et al.*, (1996) found that the proxy's knowledge and attitudes contributed a large proportion to their pain reports whereas this had little bearing on patient's reports of pain. The reliability of patient and proxy responses can also be compared. For example, Sterkenburg *et al.*, (1996) found both patient and proxy reliability was high over a three hour time period. Others have reported, in general, that patients and proxies reliability estimates are similar (Sneeuw *et al.*, 1997a; 1997b; 1998).

The vast majority of the work to date has used prospective research methods (Table 2.1). These studies do not address the validity of proxies' responses during the bereavement period. Few studies have addressed this issue as Table 2.2 illustrates. Of the four studies identified, only two studies have compared the proxy's report prospectively with those retrospectively after the patient's death, in an attempt to judge the effects of bereavement on proxy reports (Higginson *et al.*, 1994; Hinton, 1996). All the studies have made comparisons between the patient before death and with the proxy at varying times during bereavement. Consequently, assertions about the validity of the retrospective approach have had to rely to a large extent on evidence from validity studies conducted prospectively. Having examined methods used to assess the validity of proxies at the end of life, the review will discuss the findings from the validity research (Tables 2.1 and 2.2).

**Table 2.1: Prospective studies comparing patients and their significant other's ratings on aspects of the patients' experiences relevant to quality of care at the end of life.**

| Author, year, country          | design, samples   | Main findings from patient/proxy comparison   |
|--------------------------------|---|---|
| McCusker & Stoddard, 1984; USA | Repeated measures; 66 chronically or terminally ill patients; carers            | At entry: similar mean scores for physical and psychosocial aspects of the perceived health status; recreations/pastimes, rated higher by patients; sleep/rest. rated higher by proxies; 3 months f/u: agreement similar to those at entry; agreement lower if the patient had a terminal illness, lower education, not living with the proxy |
| Curtis & Fernsler, 1989; USA   | Cross-sectional; 23 terminally ill patients receiving home hospice care; carers | No differences for QL and symptoms, functioning, life and sexual satisfaction, fun, medical costs and usefulness; proxies reported more pain  |

**Table 2.1: Prospective studies comparing patients and their significant other's ratings on aspects of the patients' experiences relevant to quality of care at the end of life.**

| Author, year, country              | design, samples  | Main findings from patient/proxy comparison   |
|------------------------------------|--|---|
| O'Brien & Francis, 1988; USA       | Cross-sectional; 42 cancer patients; 50% with stage III distant spread; next-of-kin                    | Agreement good for presence of pain, cancer and treatment related questions less agreement for pain frequency and duration of pain; agreement poor for intensity and effects on activities; proxies rated pain as more intense; agreement related to the proxy living with the patient and familiarity with the patient's pain  |
| Farrow & Samet, 1990; USA          | Cross-sectional; 622 elderly cancer patients; SO   | Agreement good for factual questions regarding medical history, support and life events; less agreement on attitudes and personal information regarding support and life events. ADL agreement ranged from fair to good   |
| Higginson <i>et al.</i> , 1990; UK | Repeated measures; cancer patients receiving palliative care; 65 at entry; 12 at 4-6 weeks f/u; carers | At entry: Patient anxiety pain and symptom control rated higher by proxies; non-significant tendency for proxies to rate family anxiety higher; proxies rated more problems; tendency for proxies to be more satisfied with services4-6 weeks f/u: similar to entry; in one area patients' and proxies' ratings of general practitioner and district nurse improved                               |
| Clipp & George 1992; USA           | Cross-sectional; 30 patients with active meta-static disease; carers                                   | Agreement good for "objective" ADL and duration of caring; moderate agreement for average daily pain, current and total physical symptoms, feeling depressed and marital distress; agreement poor for "subjective" aspects, such as thoughts and feelings about the illness and worst pain in the month; proxies perceptions of illness experience more negative                                  |
| Dar <i>et al.</i> , 1992; Israel   | Cross-sectional; 40 meta-static cancer patients; spouses   | Agreement good for worse pain, pain levels; proxy's rated average level slightly higher; patients underestimated the degree of spousal distress caused by pain; proxies less satisfied with the amount of support coping with pain; proxies more concerned about cancer pain, less positive about pain relief and duration of relief; patient's stoic attitudes related to proxies pain estimates |
| Higginson & McCarthy, 1993; UK     | Cross-sectional; 67 cancer patients receiving palliative care; family                                  | Proxies' ratings higher than patients for patient anxiety, pain and symptom control; higher overall score; no significant differences for family anxiety, practical aid, wasted time communication with health professionals.   |
| Spiller & Alexander, 1993; UK      | Cross-sectional; 18 terminally ill patients; carers  | No differences in the frequency of symptoms; lack of agreement for physical and emotional symptoms except for sickness; agreement good for adequacy of support and preferred place of death; proxies viewed the patient's outlook and emotional state more negatively; significant differences in ratings of anxiety  |
| Blazeby <i>et al.</i> , 1995; UK   | Cross-sectional; 39 oesophageal cancer patients; 41% palliation treatment; carers                      | Levels of agreement were moderate for pain, emesis, dyspnoea, diarrhoea and appetite loss; fair agreement for financial impact, fatigue and emotional functioning; agreement poorest for dysphagia  |

**Table 2.1: Prospective studies comparing patients and their significant other's ratings on aspects of the patients' experiences relevant to quality of care at the end of life.**

| Author, year, country                       | design, samples  | Main findings from patient/proxy comparison   |
|---|--|---|
| Field <i>et al.</i> , 1995; UK              | Cross-sectional; 28 terminally ill patients; carers  | Agreement good for help required, ADL, evaluations of care and absence/presence of physical symptoms; less agreement for symptoms of a psychological nature, symptom distress, mood, anxiety and depression; proxies reported the presence of depression more often   |
| Madison & Wilkie, 1995; USA                 | Cross-sectional; 18 lung cancer patients; SO   | Assessed on two days; understanding of pain location good; poor for pain intensity, quality and pattern; overestimated patient pain coping strategies; proxies overestimated pain intensity in patients with low levels of anxiety or internal locus of control and underestimated pain intensity in patients with high levels of anxiety or external locus of control on day 1 but not day 2 |
| Yeager <i>et al.</i> , 1995; USA            | Cross-sectional; 86 oncology outpatients with cancer-related pain; 70% palliation or control treatment; carers | No differences pain knowledge; patients were more likely to disagree with the statement “pain medicines often interfere with breathing”; proxies reported higher levels of pain and greater distress for patients; patients underestimated the carers own distress from pain  |
| Elliott <i>et al.</i> , 1996; USA           | Cross-sectional; 122 cancer patients; family members   | Proxies rated pain higher and overestimated patients functional impairment; proxies knowledge and attitudes toward pain related to their pain reports; proxies who expected pain with cancer reported more pain, those who believed that the pain was manageable reported less pain   |
| Grassi <i>et al.</i> , 1996; Italy          | Repeated measures; 49 advanced cancer patients; carers   | At entry: proxies and patients concordant for overall QL; lowest agreement for perception of health and support; moderate to strong correlation for physical activity, daily living and outlook on life; 4 weeks follow up: similar findings; tendency for proxies ratings to improve over time   |
| Kurtz <i>et al.</i> , 1996; USA             | Cross-sectional; 216 cancer patients; carers   | Presence/absence of symptoms highest for fatigue, lowest for insomnia; higher percentage agreement for all symptoms from female proxies; agreement varied depending on symptom and proxy's age; proxies who over-reported symptoms reported greater impact on their physical well-being; proxies who under-reported symptoms were more optimistic and patients more depressed                 |
| McMillian, 1996; USA                        | Repeated measures; 118 terminally ill patients receiving hospice home care; caregivers; 48 at 3 weeks f/u      | At entry: moderate correlation for QL; “worst pain” mean scores similar, moderate correlation; proxies mean scores on pain relief substantially lower, weak correlation; 3 weeks f/u: moderate correlation for QL; “worst pain” mean scores similar to patients but weak correlation; mean scores on pain relief substantially lower than patients, weak correlation                          |
| Sigurdardottir <i>et al.</i> , 1996; Sweden | Cross-sectional; 43 advanced malignant melanoma; 34 family   | No differences for physical, role, emotional, social QL subscales; highest for physical functioning; lowest for social functioning; moderate to strong correlations for fatigue, QL, neurological symptoms and pain   |
| Sterkenburg <i>et al.</i> , 1996; Canada    | Repeated measures; 73 palliative care patients; 52 family  | At entry: agreement moderate for QL; 3 hours after entry: moderate agreement for QL; reliability across the two time periods high for patients and proxies  |

**Table 2.1: Prospective studies comparing patients and their significant other's ratings on aspects of the patients' experiences relevant to quality of care at the end of life.**

| Author, year, country                         | design, samples  | Main findings from patient/proxy comparison  |
|---|--|--|
| Lobchuk <i>et al.</i> , 1997; Canada          | Cross-sectional; 37 lung cancer patients receiving palliative care; carers                                 | Percentage of severe distress similar; moderately agreement for fatigue, appetite, pain frequency, cough, nausea frequency, insomnia and breathing; more agreement for distressing symptoms, except appetite and nausea frequency; agreement poor for insomnia, outlook and appearance; global symptom distress correlated; time since diagnosis associated with discrepancy   |
| Miaskowski <i>et al.</i> , 1997; USA          | Cross-sectional; 78 patients cancer-related pain; carers   | 23 patient/proxy dyads C in assessing pain intensity; in non-congruent NC dyads proxies underestimated (14) or overestimated (41); comparing C and NC dyads, NC patients had higher least pain scores, more days per week in significant pain, more anger and fatigue and lower QL; NC proxies reported greater caregiver strain   |
| Sneeuw <i>et al.</i> , 1997a; UK/ USA         | Repeated measures; brain cancer patients; 103 at entry; 87 at 1 week f/u; 89 at 4 weeks f/u; SO            | At entry: agreement moderate to good for functioning scales, 11 symptoms, financial impact, communication deficits, future uncertainty; agreement poor to fair for 6 symptoms (e.g. pain, nausea/vomiting) and bother of hair loss; proxy's reported lower levels of functioning and greater degree of fatigue and motor dysfunction, degree of bias small except for fatigue; discrepancies varied depending on the patient's physical and neurological condition; 1 week f/u: test-retest reliability moderate to good; 4 weeks f/u: similar to entry; increase in discrepancies in patients whose physical or neurological condition had deteriorated |
| Sneeuw <i>et al.</i> , 1997b; The Netherlands | Repeated measures; cancer patients receiving chemotherapy; 295 at entry; 189 at 3 month f/u; SO            | At entry: moderate agreement for QL domains; proxies reported more impaired functioning, lower levels of well-being, differences of small magnitude; less agreement on feelings and social activities for more impaired patients; more agreement for physical fitness and QL for more impaired; At 3 month f/u: agreement higher for feelings, social activities and QL; no differences for patients with good or poor performance status; proxies as efficient as patients at detecting changes over time for feeling, overall health and QL less efficient for pain, physical, daily and social activities   |
| Kristjanson <i>et al.</i> , 1998; Australia   | Cross-sectional; 78 terminally ill cancer patients; carers   | Correlation between symptom distress scores; of those inaccurate 37% overestimated and 2% underestimated; across the symptoms agreement moderate to fair, lowest levels for outlook, nausea intensity; fatigue and pain intensity; congruence related to marital status of the patient, patient gender, age of proxy, carers relationship to the family and family income; treatments related to some items of symptom distress  |
| Sneeuw <i>et al.</i> , 1998; The Netherlands  | Repeated measures; cancer patients; 307 at entry; 232 at 3 months f/u; SO                                  | At entry: moderate to good agreement; proxies reported lower total QL, physical, emotional, role and social functioning; proxies reported more fatigue, nausea, vomiting, fatigue, insomnia and pain; degree of bias small to moderate; most disagreement at moderate levels of QL; proxy/patient characteristics accounted a small amount to the level of agreement; 3 months follow up: similar to entry; patient and proxy responsive to changes over time; proxies better at detecting changes in physical and role functioning  |
| Nekolaichuk <i>et al.</i> , 1999b; Canada     | Repeated measures; 32 advanced cancer patients admitted to a tertiary palliative care unit; family members | At entry and 24 hours later; ratings of patient pain, anxiety, drowsiness, appetite, well-being, shortness of breath, tiredness and nausea; ratings generally consistent across symptoms and occasions; ratings from family members and health professionals not reported separately   |

**Table 2.1: Prospective studies comparing patients and their significant other's ratings on aspects of the patients' experiences relevant to quality of care at the end of life.**

| Author, year, country                        | design, samples  | Main findings from patient/proxy comparison  |
|--|--|--|
| Sneeuw <i>et al.</i> , 1999; The Netherlands | Cross-sectional; 90 cancer in-patients; SO   | Moderate to good agreement for physical fitness, feelings, daily and social activities, health and pain; agreement fair for overall QL; proxies reported more pain, impaired levels of feelings, daily activities, poorer overall health and QL, differences moderate to small, largest differences for emotional functioning; agreement better for patients with a high or low performance status; more observable phenomenon   |
| Wilson <i>et al.</i> , 2000; Canada          | Cross-sectional; 100 patients with metastatic cancer of breast (n=71) and close relatives prostate (n=29) and partners | Proxies for breast cancer patients- mean agreement on 12 out of 16 domains good except for emotional functioning and pain; overestimated pain, fatigue and degree of emotional disturbance; agreement between pairs from 13-72%. Proxies for prostate cancer patients- mean agreement on 14 out of 17 domains good; tended to rate more hot flushes, lower mood and the perceived importance ability to have an erection; agreement between pairs from 7-86%<br>Overall proxies better for more concrete, observable symptoms  |
| Lin, 2001; Taiwan, Republic of China         | Cross-sectional; 89 oncology in-patients with cancer related pain; carers  | No differences for pain ratings and interference; “least pain” rated higher by proxies; NC dyads showed NC proxies older, less educated; NC patients had more pain, pain interference, poorer performance status and more concerned with pain and analgesics; patients age, disease progression and beliefs were related to under/over estimations of pain   |
| Sneeuw <i>et al.</i> , 2001; The Netherlands | Cross-sectional; 72 patients with metastatic prostate cancer; spouses  | At entry and 24 hour test retest reliability good except for diarrhoea and sexual satisfaction; no differences in mean QL scores; proxies rated greater impairment for physical and role functioning, sleep disturbance and weight loss; lower sexual satisfaction; magnitude of bias slight to moderate except for sexual functioning and sexual satisfaction which was moderate to high  |
| Lobchuk & Degner, 2002a; Canada              | Cross-sectional; 98 advanced cancer patients; NOK  | High and low prevalent physical symptoms substantial to fair agreement on prevalence, fair to substantial agreement on frequency, severity, symptom distress and symptom experience for high and low prevalence physical symptoms except for dry mouth (underestimated), magnitude of bias small; psychological symptoms fair to moderate agreement for presence, tendency to overestimate number of psychological symptoms; fair to moderate agreement for psychological symptoms, significant differences in mean score on symptoms except feeling irritable, magnitude of bias small to moderate; agreement for all symptoms better on symptom dimension rather than broad questions on symptom prevalence; symptom distress over-reported compared to frequency and severity |
| Redinbaugh <i>et al.</i> , 2002; USA         | Cross-sectional; 31 patients with end-stage cancer (stage III or IV) receiving home hospice care; caregivers           | Pain intensity overestimated; accuracy of rating related to proxies who perceived patient in a lot of distress from pain, associated greater efforts at pain relief with more pain and were distressed at pain, were less accurate; patient QL related to more disparity; experience of patient pain related to accuracy; pain knowledge and patient and proxy demographics not associated   |

Key: f/u= follow up; ADL= activities of daily living; SO= significant other; QL= quality of life; C= congruent; NC= non-congruent.

**Table 2.2: Prospective/retrospective studies comparing patients and their significant other's ratings the quality of care at the end of life.**

| Study;<br>country                  | Study design;<br>sample   | Main findings  |
|------------------------------------|---|--|
| Ahmedzai <i>et al.</i> , 1988; UK  | Prospective: 40 terminally ill lung cancer patients; retrospective: SO            | Retrospective: all symptoms in the last month rated non-significantly higher except dyspnea; poor agreement for mood; tendency for proxies to rate mood items lower; non-significant tendency for proxies to report that patients had more insight about their approaching death.  |
| Cartwright & Seale, 1990; UK       | Prospective: 34 terminally ill patients; retrospective: family                    | Congruence poor for 14 symptoms except for constipation; proxies more critical of services and reported more symptoms (e.g. sleeplessness, depression); under/over-reporting equally divided for other symptoms (e.g. pain, trouble breathing)   |
| Higginson <i>et al.</i> , 1994; UK | Prospective: 6 patients receiving palliative care 7 family; retrospective: family | Agreement good for practical aid, communication and wasted time; poor for the other symptom and pain control, patient and family anxiety; pain and symptom ratings polarised to extremes of the rating scale; patient anxiety rated as less severe; family anxiety rated as more severe  |
| Hinton, 1996;UK                    | Prospective: 77 terminally ill cancer patients; SO; retrospective: 71 SO          | Prospective: agreement moderate to good for all symptoms except anorexia and confusion; Proxies prospective/retrospective: agreement moderate to good for symptoms such as immobility confusion, dyspnea and vomiting/ nausea; fair for weakness, constipation and depression; Patients prospective/retrospective: agreement moderate to good for mobility, dyspnea, anxiety, illness awareness and acceptance; fair for malaise/fever weakness, constipation,; Retrospective: agreement generally lower; pain rated more severe; weakness, malaise, depression and carer's stress under-rated; anxiety closer to patient's rating |

Key: f/u= follow up; ADL= activities of daily living; SO= significant other; QL= quality of life; C=congruent; NC= non-congruent.

**2.4 Discussion**

**2.4.1 Levels of agreement between the patient and proxy**

In spite of the heterogeneity of the research assessing the validity of proxies' reports, some consistencies do exist across the literature. The evidence suggests that information that relies on concrete, observable phenomenon, such as service provision, service evaluation and factual information, tends to have good levels of agreement (O'Brien & Francis, 1988; Farrow & Samet, 1990; Clipp & George, 1992; Higginson & McCarthy, 1993; Spiller & Alexander, 1993; Field *et al.*, 1995). This is further evident in retrospective reports of the practical help needed and the depth of information given to patients and their families from health professionals (Higginson *et al.*, 1994). However, there are some exceptions with greater satisfaction reported by proxies prospectively (Higginson *et al.*, 1990). One explanation for this might be relative's reluctance to criticise services (Glickman, 1997).

On the whole, agreement is better for those symptoms and aspects of the patient's functioning that are more overt, for example immobility, activities of daily living, and for some

symptoms such as fatigue, dyspnea and vomiting (McCusker & Stoddard, 1984; Farrow & Samet, 1990; Clipp & George, 1992; Spiller & Alexander, 1993; Higginson *et al.*, 1994; Blazeby, Williams *et al.*, 1995; Field *et al.*, 1995; Grassi *et al.*, 1996; Hinton, 1996; Kurtz *et al.*, 1996; Sigurdardottir *et al.*, 1996; Lobchuk, Kristjanson *et al.*, 1997a; 1997b; Kristjanson *et al.*, 1998; Sneeuw *et al.*, 1998; Sneeuw *et al.*, 1999). This finding is mirrored in Hinton's (1996) prospective and retrospective comparison. Here, proxies can draw on cues such as non-verbal behaviour when judging the patients' experiences.

With few exceptions, agreement is poorest for aspects of the patient's experience that are more subjective in nature like the patients' pain, feelings and thoughts (Ahmedzai *et al.*, 1988; Farrow & Samet, 1990; Higginson *et al.*, 1990; Clipp & George, 1992; Higginson & McCarthy, 1993; Spiller & Alexander, 1993; Higginson *et al.*, 1994; Blazeby *et al.*, 1995; Field *et al.*, 1995; Hinton, 1996; Lobchuk *et al.*, 1997; Sneeuw *et al.*, 1997a; Wilson *et al.*, 2000; Lobchuk & Degner, 2002a). Comparing patient-proxy reports for these aspects at entry to the study and three months later, has shown that agreement improves over time (Sneeuw *et al.*, 1997b). Examining proxies' responsiveness to changes in the patient's condition over time, one study found that proxies were better at detecting changes in more objective aspects of the patients' experiences such as physical and role functioning (Sneeuw *et al.*, 1998). In contrast, Sneeuw *et al.* (1997b) found that proxies were less efficient for physical fitness, daily activities and also for social activities and pain.

At the group level, systematic biases were evident for pain, with the majority of studies finding that pain was viewed as more frequent and severe by proxies (O'Brien & Francis, 1988; Higginson *et al.*, 1990; Higginson & McCarthy, 1993; Madison & Wilkie, 1995; Yeager *et al.*, 1995; Elliott *et al.*, 1996; Sneeuw *et al.*, 1998; 1999; Wilson *et al.*, 2000; Redinbaugh *et al.*, 2002). In studies where proxies underestimated patient's pain the proportions of proxies underestimating were small (Miaskowski *et al.*, 1997; Linn, 2001). This contrasts to the findings of health professional who tend to underestimate patients' pain (Grossman *et al.*, 1991; Higginson & McCarthy, 1993; Au *et al.*, 1994; Higginson & McCarthy, 1993; Nekolaichuk *et al.*, 1999a; Wilson *et al.*, 2000). When asked prospectively to report on mood, particularly anxiety and depression, proxies were also shown to overestimate the severity of mood disturbance (Higginson *et al.*, 1990; Higginson & McCarthy, 1993; Spiller & Alexander, 1993; Field *et al.*, 1995; Sneeuw *et al.*, 1997b; 1999; Wilson *et al.*, 2000; Lobchuk & Degner, 2002a). Likewise, symptoms were rated as more distressing by proxies (Clipp & George, 1992; Madison & Wilkie, 1995; Yeager *et al.*, 1995; Lobchuk *et al.*, 1997; Redinbaugh *et al.*, 2002). The evidence was less conclusive in finding a systematic bias for proxies' reports of patients' levels of functioning. Some studies indicated that proxies overestimated patient's impairment compared to the patient's own estimates (Elliott *et al.*, 1996; Sneeuw *et al.*,

1997a; 1997b; 1998; 1999), whereas others have found no such bias (Farrow & Samet, 1990; Clipp & George, 1992; Field *et al.*, 1995; Grassi *et al.*, 1996; Sigurdardottir *et al.*, 1996). Studies that have calculated the magnitude of the bias have shown that it tends to be moderate to small, with more substantial bias for feelings, emotions and fatigue (Sneeuw *et al.*, 1997a; 1997b; 1998; 1999; 2001; Lobchuk & Degner, 2002a).

Findings from studies that have assessed proxies' reports before and after the patient's death vary from prospective studies. Overall, proxies' prospective ratings are more in agreement with the patients' reports than their retrospective ratings. Higginson *et al.* (1994) found that proxies' reports of pain diverged from their prospective reports, becoming polarised to the extremes of the rating scale. Though, the study sample was small. This change in report over the terminal and bereavement phases is supported to some extent by Hinton's (1996) findings. Here, pain ratings were rated more severely retrospectively when compared to the patients' and proxies' prospective ratings. In contrast, depression was under-rated retrospectively by proxies (Ahmedzai *et al.*, 1988; Hinton, 1996). This offers support to the assertion that proxies' accounts may be a reflection of their own thoughts and feelings during bereavement (Higginson *et al.*, 1994). Yet, the findings for mood contrast to those in the prospective validity literature. After death reports by proxies on patient's anxiety revealed that anxiety was rated as less severe and, in some instances, was more in line with the patients' own ratings than their own prospective reports (Ahmedzai *et al.*, 1988; Higginson *et al.*, 1994; Hinton, 1996). This suggests that, for some aspects at least, proxies may be able to view the patients' experiences more objectively from a distance after the events. Or in the case of depression proxies may minimise the degree of patients' psychological distress retrospectively. The following sections will explore factors associated with the validity of proxies' reports and offer possible explanations for the findings.

## **2.4.2 Factors associated with the patient and proxy**

### ***Patient and proxy demographics***

Evidence for a relationship between levels of agreement between patients and proxies, and demographic information such as age, ethnicity, gender, socio-economic status and educational level have been inconsistent. Of those studies that have found that demographics do have a bearing on agreement, their influence appears to be mixed (O'Brien & Francis, 1988) or a function of the symptom evaluated (Kristjanson *et al.*, 1998). For example, older proxies were less accurate in reporting pain frequency than younger proxies (Kristjanson *et al.*, 1998). Two studies indicate that female proxies may be more congruent in their reports than male proxies, although the effect tends to be limited (Miaskowski *et al.*, 1997; Kristjanson *et al.*, 1998). Only one study found that agreement was lower if the patient had a lower educational



standard (McCusker & Stoddard, 1984). Few definitive conclusions can be drawn from the array of findings.

### ***Level of patient's impairment***

It has been argued that the patient's caregiver may be more aware of events than patients in cases where the patient is tired or confused, and so better able to report on the patient's experience (Cartwright & Seale, 1990). This could account for lack of agreement between the patient and proxy. There is evidence to support this, Miaskowski *et al.* (1997) found that patients in non-congruent patient-proxy dyads were more likely to have a lower quality of life, more fatigue and increased levels of confusion and mood disturbance. Comparing reports by chronically ill patients and their caregivers using the Sickness Impact Profile, McCusker and Stoddard (1984) found that levels of agreement were lower if the patient had a terminal illness. However, the influence of the patient's condition on levels of agreement does not appear to be consistent. Patient's cognitive, neurological or psychiatric status has been shown not to affect agreement (McCusker & Stoddard, 1984; Farrow & Samet, 1990). Discrepancies across the studies may be due to differences in assessments and in the extent of the physical and mental impairment in the patients studied. Further to this, the findings have been difficult to explain in light of studies using repeated measures designs. Sneeuw *et al.* (1997a) in a study of patients with brain cancer found that agreement on quality of life was lower for patients with poorer functioning, cognitive impairment, minor confusion at entry to the study and at one month follow up. Similar results were found by Sneeuw *et al.* (1997b) where agreement was lower for feelings and social activities, but agreement was better for physical functioning and quality of life when patients were more impaired. However, these findings did not persist over time. In other studies Sneeuw *et al.* (1998; 1999) have found a curvilinear relationship with patients at moderate or mild levels of impairment with the greatest number of discrepancies and lowest agreement at entry and at follow up.

### ***The patient and proxy relationship***

The spouse was predominately the significant other or caregiver in the studies reviewed (see Table 2.1 and Table 2.2). Therefore, it is difficult to identify who is the most appropriate person to act as the patient's proxy. Some studies found that agreement was not dependant on the family member being the spouse (McCusker & Stoddard, 1984; Higginson *et al.*, 1994), while others found that spouses were more congruent than other family members (Miaskowski *et al.*, 1997), or were better for some aspects of the patients' experiences such as health status, life events and social network (Farrow & Samet, 1990; Kristjanson *et al.*, 1998). These findings are in line with those from health research and epidemiological surveys where spouses are considered best able to report on the patient's experiences (Moore *et al.*, 1970; Okamoto,

2000; Novella, *et al.*, 2001). Differences across the studies may be due to the number of opportunities the proxy has to observe and interact with the patient during the illness. Indeed, proxies living with the patient have been shown to have higher levels of agreement (McCusker & Stoddard, 1984; O'Brien & Francis, 1988). Nearing death agreement is also higher between proxies where the patient died at home (Higginson *et al.*, 1994). Given these findings, it is reasonable to assume that those closest to the patient, in many cases the spouse, are better informed as they share their partner's experiences. Thus it is not the spouse *per se*, but the relationship between the proxy and patient that is most important. Any factors that affect the relationship, particularly ones that influence the ability of the patient and proxy to share their thoughts and feelings may diminish the proxies' ability to accurately reflect the patient's experience (Clipp & George, 1992).

### ***Caregiver burden***

McWhinney (1989) suggests that family caregivers and patients may have difficulties in sharing their feelings. The perceived burden of caring does appear to influence proxy caregiver's ability to accurately report on aspects of the patient's experience (Kurtz *et al.*, 1996; Sneeuw *et al.*, 1998). The findings concur with assessment of quality of life in other groups (discussed in Section 1.4.2). Sneeuw *et al.* (1998) found that the intensity of the caregiving role was related to proxies' ratings of the patient's quality of life. Non-congruent patient-proxy dyads reported greater caregiver strain than congruent dyads when assessing patient's pain (Miaskowski *et al.*, 1997). Clipp and George (1992) found that differences between patients and proxies accounts of depression, pain, functioning and coping might be attributable to the patient not sharing their feelings, as patients in the study tended to view their marriage more negatively than did spouses. The burden of caring might influence proxies' perceptions of events, as proxies' beliefs regarding the patient's suffering create an impression that things are worse than they really are (Lobchuk *et al.*, 1997). Thus, the views of caregivers may be biased on account of their own concerns and expectations regarding their role. The actual burden of caring occurs before the patient's death during the terminal phases of the illness. Prospectively agreement on patients' quality of life is affected by the proxies' own health and quality of life, with poorer levels predictive of the difference between the patient and proxy (Sneeuw *et al.*, 1998). What is not known is how caregiver burden effects proxies' responses after the patient's death. The negative physical and psychological impact of the perceived burden of caring for patients during advanced illness is well documented (Stez, 1987). It is likely that the negative consequences of caring transcend into bereavement. Kelly *et al.* (1999), in a longitudinal analysis, found that the level of psychological symptoms, lower levels of practi-

cal assistance, the quality of the relationship and coping style in the terminal phase of the illness were among factors predictive of poorer bereavement outcomes.

However, all studies to date, that have examined the role of caregiving on levels of patient/proxy agreement have focused on the negative effects of caring, conceptualising caring as a negative experience. This has been to the detriment of fully exploring the multifaceted role of caring with respect to congruency. For example, whether the positive aspects of caring, identified by researchers in the area of caregiving (Nolan *et al.*, 1994; Grant *et al.*, 1998), influence the way in which proxies perceive patients' symptoms. Future work would benefit from a comprehensive exploration of this area, particularly in light of the numbers of proxies providing end-of-life care.

### ***Expectations and beliefs***

The way in which events are perceived is based on the individual's frame of reference. That is, the context within which they are embedded, their experiences, attitudes and beliefs (Vernon, 1970). Therefore, it is not surprising to find that there is incomplete agreement given that the patient and the proxy have their own perspectives of events. For example, symptoms such as depression, anxiety and pain may be interpreted as a normal part of the illness experience, whereas for others these symptoms may be viewed as unacceptable and worth reporting. Differences in reporting these symptoms may occur if there is a mismatch in the perceptions of the patient and proxy. Elliott *et al.* (1996) demonstrated that family members' report of patients' pain is strongly related to their attitudes and knowledge. Proxies, who expected pain with cancer reported more pain, whereas those who believed cancer pain management had no ceiling dose for opioids and were not concerned about addiction, reported that patients had less pain. In addition to these considerations, proxies hold beliefs about the patient and how they respond to the experience based on their shared experiences. These beliefs may influence the perception of events or proxies may take into consideration particular beliefs they hold about the patient when making judgements. For example, stoicism in patients was associated with spouses' overestimation of pain (Dar *et al.*, 1992). This is consistent with the literature in social psychology where judgements by others are based on assumptions regarding the individual's personality characteristics (Sudman *et al.*, 1996). Patients own beliefs can affect the validity of their responses and in turn lead to poor agreement. Sneeuw *et al.* (1998) found that patient's tendency toward socially desirable responses was a predictor of patient and proxy agreement.

### ***Coping***

How patients cope with their illness prior to death and how proxies cope during the terminal and bereavement phases could have implications for the validity of their responses. For exam-

ple, discrepancies may result if patients conceal their feelings to protect their significant other. Dar *et al.* (1992) found that 60% of patients admitted hiding their pain so as not to upset their spouses. The reverse may be true for proxies as patients have been found to underestimate the degree of distress caused to proxies (Dar *et al.*, 1992; Yeager *et al.*, 1995). Disruptions in communication are not uncommon as patients and their partners try to cope with the patient's illness and approaching death (Baider & Sarell, 1984). Patients' feelings and less visible symptoms, such as mood states, may be easier to conceal than more overt symptoms such as vomiting and breathlessness. Coping strategies such as denial may inhibit the open expression of feelings between the patient and proxy. As identified in Section 1.4, denial by the patient has been shown to skew health professionals' reports when measuring affective states (Jennings & Muhlenkamp, 1981). Spiller and Alexander (1993), after assessing eighteen terminally ill patients and their family caregivers, found that patients had a more positive outlook and emotional state than their proxies reported. The authors concluded that differences between the ratings might be due to denial or the proxy's projection of their own feelings onto the patient. Evidence for this comes from Kurtz *et al.* (1996) who found that in patient-proxy dyads where the patient identified a symptom but the proxy did not, that the patients tended to be more depressed and the proxy more optimistic. They suggest that the depressed patient may over-report symptoms or the proxy may minimise and under-report the number of symptoms. Miaskowski *et al.* (1997) also found that proxies underestimated pain intensity in more depressed patients. Therefore, lack of agreement may be a function of how both the patient and proxy interact and how they cope with events. Once again, how the proxy copes with events before the patient's death may carry over into bereavement and influence their retrospective responses.

### **2.4.3 Factors associated with assessment**

#### ***Assessment instrument***

Responding to a question is a complex cognitive process, particularly if it requires making judgements regarding other people's experiences. It consists not only of attending, perceiving and encoding information pertaining to an event, but also recalling material and judging its appropriateness in relation to the question and the response format. Errors in responding can occur at any stage. Thus, some degree of inaccuracy in proxies' reports can be attributed to the instrument itself. As Sudman and Braburn (1974) point out, variables associated with the task itself are an important source of error. The cognitive psychology and survey literature is abounding with potential biasing factors, amongst which are question length, ordering of questions, mode of administration, response options and content (Jobe *et al.*, 1993). The findings from the studies concur with this literature and have been identified in other studies

examining the validity of proxies (Pickle *et al.*, 1983; Magaziner *et al.*, 1988; 1996; 1997). Agreement between cancer patients and their next-of-kin proxies was found to deteriorate when the questions became more complicated and the number of response options increased (O'Brien & Francis, 1988). Further evidence comes from Sneeuw *et al.* (1997a), where the level of exact patient and proxy agreement was a function of the response format, with high agreement for dichotomous options and low agreement for two items with seven point scales. Questions and response options that are vague and ambiguous may also compromise validity, as the proxy may misinterpret the meaning of the question. If the task is too cognitively demanding, proxies may simply choose an answer that appears satisfactory. This is known as “*satisficing*” and can result in misleading information (Krosnick, 1991). One can postulate that if the task is too emotionally demanding, then satisficing may also occur as individuals seek to reduce the emotional strain of actively searching their memories. A mediating factor that can improve recall is the participant’s level of motivation. The higher it is, the more likely they are to do a thorough search through memory thus improving accuracy of recall (Sudman *et al.*, 1996).

### ***Content of the assessment***

As identified earlier and consistent with the literature assessing proxies in other areas of health care, proxies are better able to report on more objective and visible aspects of the patients’ experiences. It has been suggested that this occurs because the difficulty of a reporting task is determined by its content (Hughes & Preski, 1997). Thus, questions that assess more abstract or subjective phenomenon are more likely to have recall errors because they require inferences. Since the proxy is not experiencing the phenomenon themselves, they have to rely on less visible cues. Confusion may occur when interpreting these cues because some symptoms, such as depression, may overlap with other somatic symptoms (Brugha, 1993), or the proxy may be unable to identify the cues through lack of understanding of the nature and cause of the symptom. In such instances, the experience may be misidentified. Such arguments provide an explanation for the inconsistent findings for overall quality of life where agreement ranges from fair to good (Grassi *et al.*, 1996; McMillan, 1996; Sigurdardottir *et al.*, 1996; Sterkenburg *et al.*, 1996; Sneeuw *et al.*, 1997a: 1997b; 1998; 1999; Wilson *et al.*, 2000; Sneeuw *et al.*, 2001; Redinbaugh *et al.*, 2002). Since quality of life is a multidimensional construct incorporating subjective and objective elements. The same can be said for pain assessment where affective components of pain are less congruent with the patient’s reports than physical components (Dar *et al.*, 1992; Madison & Wilkie, 1995).

The content of the material being assessed is a significant factor in recall. For example, the capacity to recall the specifics of regular, mundane activities, such as food intake, deterio-

rates rapidly (Armstrong *et al.*, 2000). By comparison, memory for infrequent exceptional events, like health care visits and hospitalisation, declines less rapidly (Mathiowetz, 1988). The frequency and saliency of an event is highly relevant to retrospective interviews because proxies are likely to be asked to report on the frequency and severity of patients' symptoms. Although not investigated in the studies (Table 2.1 & 2.2), research suggests that memory for these aspects can be selectively biased. Events that are highly salient, that stand out, tend to be remembered (Bower *et al.*, 1979). In addition to this, salient events may be over-reported (Huber & Power, 1985). This may account for the tendency of relatives to overestimate patients' levels of pain, as pain relief is often ranked as their highest priority (Kristjanson, 1989). Less salient events may be prone to inaccuracies, as they become typical and less distinct. As a result more easily forgotten or merged in to a generic memory that is a summary of events (Linton, 1982). In addition to this, research into eyewitness testimony, both in field and laboratory settings, has demonstrated that highly emotional events are recalled better than neutral ones (Christianson, 1992). Since the death of a significant other is an important and rare event these findings suggest that memory should be better than for more usual life events. However, recall of specific aspects of the patients' experiences may be inaccurate because of their occurrence and significance.

### ***Assessment period***

The actual period being assessed varies enormously across the validity studies from the past few hours (Linn, 2001) up to the last year of life (Cartwright & Seale, 1990). This has made it difficult to examine what influence the assessment period has on the validity of proxies' responses even though it is likely to affect the strategy proxies use to recall information. For example, it is unlikely that assessments over a year will result in a specific search for each episode of an event. In these instances, individuals may resort to strategies that approximate events rather than a precise recall of each episode (Linton, 1982). Although this may appear problematic, it does not necessarily undermine the validity of proxies' reports, as retrospective surveys may not necessarily require specific information. Instead, a general idea about the adequacy of care may be all that is required.

### ***Timing of the assessment***

From the research on memory it has been shown that one factor important in remembering is the time interval between encoding and retrieval (Baddley, 1990). Consequently, the shorter the intervening period, the easier it should be to recall information retrospectively. This may explain why agreement between patients and proxies is better prospectively when compared to retrospective reports (Hinton, 1996); the recall of events are in close temporal relation to the actual event and the encoding of the material into memory. Thus, recall error is more

likely to be problematic retrospectively where proxies are often expected to remember events several months after the patient's death. Due to the sensitive nature of the information, the timing of the interview also has ethical implications. It may be morally inappropriate to interview too soon after the patient's death.

Differences in the timing of retrospective assessments have made it difficult to ascertain the most appropriate time to interview proxies after the patient's death. Responses by proxies to questions during bereavement have been shown to diverge from their earlier recollections in the terminal phase of the patient's illness (Higginson *et al.* 1994; Hinton, 1996). Similarly, responses regarding patients' symptoms, care received and need for help have been shown to change between interviews at 3 and 9 months during bereavement (Cartwright *et al.*, 1973). More research is needed to investigate how proxies' reports change during bereavement and whether these changes are significant enough to question the validity of proxies' reports. The findings concur with the memory literature, which has found that later recollections may be inconsistent with earlier recollections and that informants may not realise that their recollections have varied over time (Ross, 1989). This is also consistent with the notion that memory is constructed and reconstructed (Conway, 1992). As these constructions are embedded in the individuals general knowledge about a particular event, they may be influenced by the proxies' own emotions.

### ***Memory and emotions***

Emotions are an integral part of attention, perception and retrieval of information, and yet the effect of bereavement on the accuracy of proxies' memories for events has not been explored. Notwithstanding the inherent challenges of this type of research, it is likely because it is difficult to identify which of the many factors associated with grief affect the memories of bereaved relatives. A mixture of emotions often marks this period as they cope with the impending loss and its associated challenges. This can produce noticeable levels of stress (Keitel *et al.*, 1999) and depressed mood (Shuchter & Zisook, 1993). Research into how emotions influence memory has demonstrated that there is a mood congruent memory effect (Bower, 1981; Bower, 1987). This is described as: "enhanced coding and/or retrieval of material the affective valence of which is congruent with the ongoing mood" (Blaney, 1986, p.229). A pervasive finding is that individuals with depression or depressed mood have been found to recall more negative stimuli (information, events), which is consistent with their mood-state (Matt *et al.*, 1992; Eich, 1995). Although a similar effect has been identified for individuals with anxiety, the findings have been less consistent (Mogg *et al.*, 1987; Eysenck & Mogg, 1992).

Further to this, emotions can create attentional biases. Bower (1992) suggests that individuals pay more attention to information that is consistent with their mood state. In other words, anxious individuals tend to focus on threatening information, whereas, those who are depressed are likely to attend to sadness and loss (Mineka & Sutton, 1992). Evidence from individuals with clinical anxiety (Burke & Matthews, 1992) and non-clinical anxiety (Broadbent & Broadbent, 1988) supports the proposition that anxious individuals selectively attend to threat information. Attentional processes have also been given a central role in explaining anxiety states in theories of anxiety (e.g. Eysenck, 1992). However, the results from studies investigating attentional biases in depression have been mixed (McCabe & Gotlib, 1995; Bradley *et al.*, 1997; Williams *et al.*, 2000)

The nature of the event itself can also influence attention. Research from cognitive psychology suggests that emotional events that create physiological arousal are liable to attentional biases. These biases are thought to narrow attention to the central details of the event; reducing attention for peripheral aspects (Christianson, 1992). For highly emotional events, it has also been found that recall for the central details are improved for several months following the event (Christianson & Loftus, 1987). These findings may explain why proxies' reports vary from those of patients, as they may be focusing on specific aspects that they believe are significant.

Emotions can also influence the strategy that individual's adopt to recollect material. These strategies vary depending on the task and the individual's cognitive abilities. As noted earlier, a generic memory that is a summary of events may be recalled rather than a search through episodic memory (Linton, 1982). This strategy can be controlled and a more specific memory strategy employed as the need arises. However, depressed individuals (Williams & Dritschel, 1988) and older people with reduced memory capacity (Winthorpe & Rabbit, 1988) are two groups who have been found to have difficulty in doing this. Furthermore, reports on other peoples' behaviour and attitudes are more likely to rely on estimation strategies as the material is not as elaborately coded in memory and therefore as accessible to recall, compared to self-reports (Sudman *et al.*, 1996). Therefore some individuals may be more likely to use estimation strategies.

The research has highlighted how fallible memory is for events and the complex relationship between memory and emotion. Judgements based on these memories are likely to be biased (Kahneman *et al.*, 1982). An example is the "*availability heuristic*", which states that the likelihood of an event is judged by the ease with which events can be recalled (Tversky & Kahneman, 1973). Moreover, emotions can influence judgement themselves (Schwarz & Clore, 1989). Despite the significance the proxy's emotional state, few studies have investigated the proxy's psychological health status directly. Where studies have investigated it,



researchers found a non-significant tendency for proxies in non-congruent patient-proxy dyads to have poorer mood states than congruent dyads (Miaskowski *et al.*, 1997). Kurtz *et al.* (1996) found caregiver depression was not a significant factor in levels of patient and proxy concordance. Clearly, this is an area that requires further investigation.

#### **2.4.4 Factors associated with the validity studies**

The research presented so far has summarised the many factors that can potentially affect the validity of proxies' reports. This section deals with the quality of the studies that have investigated this aspect.

##### ***Assessment periods***

Some discrepancies between patients and their proxies may be the result of the research itself. This is particularly striking in the retrospective/prospective studies, when investigators are not comparing like with like. As Higginson *et al.* (1994) highlights, in a study conducted by Ahmedzai *et al.* (1988), patient reports were recorded at different times before death, whereas, relatives were asked to report on the last month of life. Under these circumstances it is not surprising to find differences when patient and proxy views are compared. Cartwright and Seale (1990) acknowledge that differences could be due to the fact that interviews were conducted with patients at variable times before death. As a consequence differences might be attributable to the differences in periods assessed. Moreover, the proxy might have reported on experiences between the patient interview and death not covered by the patient interview.

##### ***Measures of association***

Problems have been identified with the statistics used to assess the measure of agreement. A number of studies have used the Pearson *r* statistic, Spearman *rho* or Kendall's *tau* (Ahmedzai *et al.*, 1988; Clipp & George, 1992; Dar *et al.*, 1992; Grassi *et al.*, 1996; Sigurdardottir *et al.*, 1996; Sneeuw *et al.*, 1998). However, these type of correlations do not indicate that one individual's rating reflects another: the relationship between two raters could be significant if the ratings were consistently higher or lower than the patient's ratings (Nelson *et al.*, 1990). An alternative statistic that measures the degree of agreement between sources of continuous data, which accounts for systematic mean differences and chance is the interclass correlation coefficient (ICC) (Landis and Koch, 1977). The equivalent for categorical or dichotomous data, allowing for chance agreement is Cohen's kappa (Cohen, 1960). Comparing statistical measures of agreement Linn (2001) found that using the ICC patient-proxy ratings of pain agreement was high in contrast the kappa coefficient, though significant, indicated poor levels of agreement. An explanation for this could have been because the kappa was unweighted, that is its value was based on one-to-one agreement between the patient and proxy. Weighting

the kappa allows for some degree of disagreement within acceptable clinical or theoretical boundaries (Cohen, 1968). This is an important consideration when evaluating the validity studies, as agreement will vary as a function of the standard chosen. For some investigators strict one-to-one agreement is necessary (Higginson *et al.*, 1990; Higginson & McCarthy, 1993; Spiller & Alexander, 1993; Hinton, 1996; Lobchuk *et al.*, 1997; Kristjanson *et al.*, 1998), while for others, agreement can vary from the patient's response to a certain extent (Higginson *et al.*, 1994; Sneeuw *et al.*, 1997a; 1997b; 1998; 1999). Of more importance when deciding on level of agreement, is the necessity for a strict agreement. Is it more important to know that the patient was in a certain amount of pain regardless of whether the response criteria were accurately matched? In such instances, weighting the level of association to allow for some level of disagreement would be more appropriate.

A related issue is that of systematic bias. Several studies have examined proxies tendency to over- or underestimate patients' experiences. However, few have actually quantified the amount of bias or attempted to explain the significance of their findings in clinically meaningful ways (Sneeuw *et al.*, 1997a; 1997b; 1998; 1999; 2001; Lobchuk & Degner, 2002a). Yet for researchers who rely on proxy reports this is requisite when interpreting the information given by proxies. Furthermore, it allows meaningful comparisons across studies (Lobchuk & Degner 2002b). One approach put forward by Sneeuw and colleagues (1998) is to standardise the mean difference by relating the scores to their standard deviations. Using Cohen's (1988) standardised differences (effect size  $d$ ) to interpret the size of the differences;  $d=.2$  (small difference),  $d=.5$  (moderate difference) and  $d=.8$  (large difference). At best this provides some means by which to interpret differences in the absence of guidelines as to what degree of difference is clinically meaningful (Sneeuw *et al.*, 1999).

### **Sample**

The sample size of some of the validity research is relatively small (Spiller & Alexander, 1993; Higginson *et al.*, 1994; Field *et al.*, 1995). For example, in a study by Higginson *et al.* (1994) only six patients and seven family members were interviewed within three weeks of the death. Comparisons of prospective and retrospective ratings on such small samples are questionable and require further investigation to substantiate their conclusions. The problem of sample size has been addressed by Sneeuw and colleagues (1997a; 1997b; 1998; 1999) where sample sizes range between 90 and 307 patients. Much of the evidence is based on convenient samples with patients who are willing and able to complete measures because of the difficulty in obtaining information from less able patients. This does limit the generalisability of their findings. As a means to overcome this problem Sneeuw *et al.* (1998) proposed examining trends as a function of the patients condition. Though insightful the results have been

inconclusive and have still relied on patients who are able to complete measures. This will remain an area of contention because of the inherent difficulties in assessing patients who are so ill.

### ***Outcome measures***

Within palliative care there are an increasing number of tools for measuring aspects of the patients experiences, relating to symptoms, and services. Whatever, the method a necessity for agreement between patient and proxy is reliable and valid measures. In a substantial proportion of studies there is a lack of evidence to suggest that this criteria has been met (Ahmedzai *et al.*, 1988; Cartwright & Seale, 1990). In Hinton's (1996) study the open-ended interview style may have been prone to subjective biases. Almost all the validity studies have used some form of interview schedule to assess palliative care outcomes, although few studies have used the same instruments. Added to this, Lobchuk and Degner (2002b) point out that lack of definition of conceptually related terms makes it difficult to interpret the findings. Examples include, nervousness, worrying and anxiety (e.g. Higginson *et al.*, 1994; Lobchuk & Degner, 2002a; Madison & Wilkie, 1995). Lack of consistency and clear definitions impedes cross study comparisons. In spite of this, some consistencies do emerge, particularly in relation to the content of the assessment. Where the same measures have been used, for example the European Organisation for Research and Treatment of Cancer quality of life instrument (EORTC QLQ-C30) comparisons are difficult because two of the studies were with disease specific patients with either oesophageal or brain cancer (Blazeby *et al.*, 1995; Sneeuw *et al.*, 1997a).

### **2.4.5 Summary and conclusions**

This review provides encouraging evidence for the use of proxies at the end of life, and shows that they can be relied upon to represent certain aspects of the patient's experience, most notably for service provision and evaluation, and for symptoms that are more observable in nature. However, care must be taken when using proxies to report on aspects of the patient's experience that are more subjective, such as pain and affective states. As this review has identified, these reports are prone to biases for various reasons (summary in Appendix A). The findings and conclusions are consistent with Sprangers and Aaronson's (1992) review and a recent review conducted by Lobchuk and Degner (2002b). The present review differs from these reviews as Sprangers and Aaronson (1992) examined both significant others and health professionals as proxies for patients with chronic disease, while Lobchuk and Degner (2002b) focused on symptoms in patients with cancer and was not specific to proxies for patients at the end of life. The focus of this paper is specifically the validity of the proxy in relation to

care at the end of life. This includes both prospective and retrospective reports and is not limited to cancer patients.

To a large extent inferences regarding the validity of retrospective proxy reports have had to be made from prospective studies due to the paucity of research on the validity of retrospective reports. It should be noted that there is evidence that there may be differences between prospective and retrospective studies. The review adds to previous work, by incorporating knowledge from other areas to help explain the research evidence and factors likely to effect validity. Most notably, from the work by psychologists and survey researchers, which has advanced our understanding of the complexities associated with responding to a question. These include the stages of attending, perceiving an event, encoding the information, and recalling the appropriate material according to response criteria. Further to this, are the differences between self-report and proxy responding. However, the research investigating the validity of proxies has primarily focused on levels of agreement to the detriment of further understanding why proxy views differ and whether and how greater agreement is possible. Further research is needed into this issue. Such investigations may help researchers understand and design measures to increase the validity of proxies' reports. Based on this review one approach would be to design measures that tap into more objective aspects of the patients' experiences.

The lack of agreement observed in the validity literature adds weight to the argument that the patient is best able to report on their experiences (Schipper & Levitt, 1985). This issue is not in contention here. The retrospective approach in palliative care does not distract from the prospective studies and the important contribution they make to research at the end of life. Instead, the retrospective approach and the use of proxies to evaluate patient's care can be seen as one solution to overcome obstacles that researcher's face when evaluating care of the dying. It therefore, has an important contribution as it provides a means to evaluate the care of patients who would otherwise not be represented. Furthermore, the research can be useful in identifying reasons for discrepancies and to develop strategies to improve congruence. With reliance on the patient's family as the main caregivers and more patients dying at home this will have important implications for the appropriate management of patients at the end of life.

## **2.5 Research aims**

Through this review an array of factors that can affect the validity of proxies' responses were identified. So far there has been little attention to investigate proxies' responses through into the bereavement phase. As the review highlighted there are significant gaps in the knowledge surrounding the effects of bereavement on proxy reports. Among these are the effects of proxies, emotions on judgments, the reliability of proxies' reports over time and the basis for judg-

ments on subjective aspects of the patient's experiences. Another notable omission, is the lack of reliable and valid instruments with which to assess palliative care retrospectively. In view of these findings, the aims of this thesis are to investigate these aspects more fully. More specifically the research sets out to investigate how proxies judge the patients' experiences after the patients' death and the effects of bereavement on their reports. The focus of the work is primarily on the more subjective symptoms such as pain, anxiety and depression. These are important symptoms to evaluate in the care of patients with terminal illnesses. In this congruency between the patient and proxy is known to be poorest. Therefore it is important to understand more about how these questions are answered so that improvements can be made. The following research questions are proposed:

### **2.5.1 Research question one**

What do proxies understand by the pain, anxiety and depression frequency and severity descriptors used in retrospective questionnaires that are used to assess palliative care?

#### ***Rationale***

VOICES (Views of Informal Carers- Evaluation of Services) (Addington-Hall *et al.*, 1998) is a retrospective semi-structured postal questionnaire used to assess palliative care. The questionnaire is a shortened version of the measure used in the Regional Study of the Care of the Dying (Addington-Hall & McCarthy, 1995a) which is based on earlier surveys (Cartwright *et al.*, 1973; Cartwright & Seale, 1990; Seale & Cartwright, 1994). It covers the last year of life and includes items on sources of informal and formal care; service utilisation and evaluation; information and communication with health care professionals; patients symptoms and restrictions, and the proxy's experience of bereavement and bereavement care. To date, there is no information on the reliability and validity of the measure. This is currently under investigation by the Department of Palliative Care and Policy at King's College, London. The aim is to compare and contrast what proxies understand by the pain, anxiety and depression frequency and severity descriptors used in VOICES (Addington-Hall *et al.*, 1998). This is a way of identifying ambiguous quantifiers in the questionnaires and as a means of improving the validity.

### **2.5.2 Research question two**

What are proxies recalling when they report on patients' levels of pain, depression and anxiety retrospectively and what cues they are using?

### ***Rationale***

Awareness of how proxies judge symptoms and what cues they use to reach decisions on patients' levels of pain, anxiety and depression will be useful in understanding possible biases. In addition to this, it will assist in the development of measures to assess end-of-life care retrospectively.

### **2.5.3 Research question three**

Do proxies' own emotions have a bearing on their reports of patient's pain, anxiety and depression?

### ***Rationale***

Proxies' responses will be interpreted in the light of their self-reported depression, anxiety and caregiver strain. This will help to establish whether their own emotions are reflected in their responses.

### **2.5.4 Research question four**

How do proxies' perceptions of patient's pain, anxiety and depression change during the bereavement period?

### ***Rationale***

By investigating the differences and similarities of the accounts given by proxies it will be possible to examine whether and how perceptions change over time. This may have a bearing on the use of retrospective measures, such as VOICES, during the bereavement period. If there are significant changes in the reports between different time periods then this suggests that the approach might not reflect events. It will also be a first stage to investigating the reliability of the VOICES instrument.

### **2.5.5 Summary**

This thesis therefore addresses the question of whether and how the congruence between the patient and proxy, on important questions of pain, anxiety and depression, can be improved. From the following investigations it will not be possible to assess the accuracy of proxies' responses in relation to patients' accounts, as there is no assessment from the patient during the terminal stage of their illness. Instead the methods proposed might help unravel some of the potentially biasing factors that can influence proxies' reports. For example, it might identify difficulties proxies' experience with particular questions and reasons for this. Furthermore, it will increase understanding regarding how proxies' evaluate patients' care, what cues to use to base their judgments. These will provide a better understanding of the limitations of

using proxies and facilitate the design of retrospective surveys used for assessing care for patients at the end of life.

# CHAPTER 3: Methods and procedures

## 3.1 Introduction

Previous investigations of the retrospective approach have used comparisons between the patient and proxy as a measure of validity. In so doing, they have identified particular areas wherein significant differences occur between the two accounts. However, studies such as these, which rely on quantitative methods and emphasise outcomes, have not been able to fully explain these discrepancies. A better understanding of how and why proxies' accounts of patients' experiences diverge is critical, not only for identifying the limitations of proxies, but also for facilitating the design of retrospective surveys used for assessing care for patients at the end of life. Since the investigations are exploratory and are aimed at the proxies' subjective meanings of the survey questions, a qualitative approach is best suited. Qualitative methods, as Miles and Huberman (1994) point out, can be particularly useful in the development of quantitative measures and to explain findings in quantitative research. Such methods have now become more widely accepted (Punch, 1998; Barbour, 1999). This chapter will discuss the methods used to address the research aims. Consideration is given to the theoretical and epistemology of the approach taken to make clear the position of the investigator. The design and procedure will then be detailed.

The research sets out to investigate how proxies judge the patients' experiences retrospectively. This entails an analysis of how proxies interpret and answer questions used in retrospective surveys. Furthermore, it aims to explore the effects of time and the proxies' own feelings on their reports of patient's pain, anxiety and depression. One method that is particularly well suited to these types of investigations is the cognitive interview. This will be discussed along with the method used to analyse the interviews.

## 3.2 Cognitive interviews

The cognitive interview consists of an in-depth face-to-face interaction between the investigator and the participant. Since they are intensive and sometimes lengthy, the number of participants can be quite small. Within these interviews, a number of techniques are used singularly or in combination to make inferences about the underlying cognitive processes used by participants. These include verbal protocols, probing, sorting tasks, rating tasks, paraphrasing, response latency and vignettes. The most commonly used technique is the verbal protocol or think aloud protocol (Ericsson and Simon, 1980; 1984). As the primary cognitive interview technique used in these investigations is the verbal protocol, the focus will be on this technique. Before discussing this in more detail, it is fundamental to understand the background to the method chosen.



The cognitive interview draws on work from several different fields, including cognitive and social psychology, and survey methodology. Its purpose is to improve the understanding of the underlying mental processes involved in answering survey questions and to provide insight when designing questionnaires (Jabine *et al.*, 1984). The theoretical underpinnings can be found within cognitive psychology in the information processing paradigm (Jobe & Loftus, 1991; Jobe & Mingay, 1991) and social psychology in attitude structure (Tourangeau, 1987; Tourangeau & Rasinski, 1988) and social information processing (Hippler *et al.*, 1987). These fields have provided a means by which to understand how participants' think when they are asked questions. This understanding is based on both theory and empirical evidence and has formed the basis for several models of the cognitive processes used in answering a survey questions (Oksenberg & Cannell, 1977; Tourangeau, 1984; Tourangeau & Rasinski, 1988; Willis *et al.*, 1991). These models tend to be sequential, but as Sudman *et al.* (1996) point out, participants may move back and forward between the different stages. The models generally follow a similar format as Willis (1991) describes:

1. The first stage involves comprehension of the question that guides retrieval of information from memory.
2. Decision/judgment processes guide the search through and extraction from memory. This includes decisions about the search process and the adequacy of the information retrieved.
3. Information relevant to the question at hand is retrieved from memory. This information may be all of the required information to answer the question, or may consist of estimations based on partial or incomplete information.
4. The information is recoded into the desired response format, or as in Tourangeau's (1984) model, the response is a selection between response alternatives.

As the models highlight, the cognitive interview is particularly useful for investigating cognitive issues such as comprehension, memory and decision making. This is important for detecting potential sources of error that can affect proxies' responses and for assessing the quality of the data provided. It can also be used to discern questions that are interpreted differently by different participants, identify ambiguities, rate task difficulty, and provide insight into how well participants are able to provide certain types of information.

### 3.3 Verbal protocols

The verbal protocol is one technique used in cognitive interviews that has been applied in the design of a number of surveys (Belson, 1981; Royston, 1989; Blair *et al.*, 1991; Menon 1993; Bickart *et al.* 1994). The majority of this work has concentrated on market research investigating the behaviour of consumers. The technique requires participants to verbalise their stream of consciousness and thoughts, while answering a question (concurrent verbal proto-

col) or after they have finished answering a question (retrospective verbal protocol). Verbal probing can also be used to extract information about particular terms used in the question. Like many other cognitive techniques, there is the underlying assumption that:

*“The psychological significance of an individual's thoughts and feelings, as well as the underlying cognitive processes, can be examined by content analyzing the individual's reported thoughts, ideas, images and feelings.”*

Cacioppo *et al.*, 1997, p.929

Thus, verbal reports can be used to infer the processes underlying responses (Ericsson & Simon, 1984). This is particularly useful when the participants are proxies, because it can be applied to examine the foundations upon which proxies base their judgments. In other words, verbal protocols can be used to establish the extent to which proxies' accounts are based on their own feelings and beliefs (with respect to the events surrounding the death of their significant other). The cues proxies' use when making decisions and how these are reached based on the available information. This goes some way to identifying what information they are able to give and under what circumstances they are able to give it.

### **3.3.1 Verbal protocol procedure**

It is disputed whether concurrent or retrospective verbal protocols are most appropriate (Blair & Burton, 1987; Bolton, 1991). Concurrent protocols are closely related to question answering and it is thought that verbalising thoughts may alter the conditions under which the question would normally be asked. By comparison, retrospective verbal protocols require the participant to recall what they were thinking, which may be prone to memory errors. According to Ericsson and Simon (1993), as long as the period between question and retrospective verbalisation is kept short, this should not affect memory for processing, as retrieval cues remain in short-term memory allow effective retrieval of the sequence of thoughts. Taking into account these recommendations, a retrospective verbal protocol was used for the present investigation. Proxies were asked to verbalise what they were thinking immediately following their response to a question. The retrospective verbalisation was kept in close temporal proximity to the actual survey question in order to aid recall of the thought processes involved. An additional consideration was the length of the task. Ericsson and Simon (1993) suggest that the actual sequence of thoughts in completing a short task, .5 - 10 seconds, is likely to be recalled with a high degree of accuracy and completeness. Given that the task of answering a question in the present study was brief, it is likely that proxies were able to recall their thoughts accurately.

Ericsson and Simon (1993) have devised standard guidelines for verbal protocols. The instructions are important, because they have implications for the type of responses obtained. To avoid changing the sequence of thoughts and to minimise socially motivated verbalisa-

tions, participants are not asked for justifications, rationalisations, descriptions or explanations. Instead, participants are asked to verbalise only those thoughts they were aware of when performing the task. Thus, the instructions to proxies followed those proposed by Ericsson and Simon (1993). For example, “*What were you thinking about as you answered that question?*” However, in conjunction with this, some verbal probing asking why questions can be effective in making explicit the particular cognitive processes used by the proxy during their verbalisations, such as, “*why did you choose severe to describe their pain?*” Other investigators attest to this technique (Willis *et al.*, 1991; DeMaio & Rothgeb, 1996). In addition to this, verbal probing is useful for focusing the participant on particular terms in the questions. In this study for examining the frequency descriptors (most of the time, often, sometimes, rarely), and severity descriptors (severe, significant, moderate, mild) available for proxies to describe the patient’s symptoms. For example, proxies were asked “*In your own words what comes to mind when you think of severe depression?*” Both types of verbal probes were used during the interviews. Although they were kept to a minimum, so as not to cause too much distress to proxies, and because extensive probing can make comparisons across interviews difficult (Willis *et al.*, 1991).

### **3.3.2 The validity of verbal protocols**

Verbal protocols are based on the assumption that individuals actually have conscious access to relevant cognitive processes. Nisbett and Wilson (1977) have challenged this assumption. They provide evidence to suggest that individuals’ cannot be relied upon to provide direct verbal reports to some types of cognitions. Willis *et al.* (1991) question the relevance of the research upon which this challenge is founded. They state that the focus of Nisbett and Wilson’s (1977) research was on why cognitive processes were carried out. In this they asked participants to identify the reasons for performing a behaviour. This contrasts with the emphasis of verbal protocols in survey methodology, where the focus is on what processes have occurred. Further refuting Nisbett and Wilson’s (1977) claim; Ericsson and Simon (1980, 1984) cite several studies where individuals have been very adept at verbalising cognitive processes. They suggest that verbal reports will be accurate under particular conditions. This is when: (1) the length of time between the task and verbalising the thought processes is kept short; (2) the information requested is descriptive rather than interpretative; (3) the information is normally available in short-term memory; (4) the information is specific.

The first two conditions have already been mentioned and addressed. It is difficult to say whether the information requested would normally be available in short-term memory without making assumptions about what people are thinking and have reasonable access to. Specific problems that might be expected given the context of the study are those pertaining

to the thoughts and feelings of recently bereaved participants. It is possible that coping mechanisms such as denial or distress may prevent proxies' accessing particular thoughts. This may or may not become apparent through the interviews. Anticipating this, codes were developed to help identify these issues. For instance, proxies' references to not wanting to think about the issue may be an indicator. This raises ethical concerns regarding the use of cognitive interviews with bereaved relatives. Therefore, ethical issues will be discussed more fully in Section 3.5.3. In response to the type of information, the fourth condition, this was kept specific to the question asked.

Other critics of verbal protocols question the extent to which individuals are able to access certain aspects of their thought processes. As Glass and Arnkoff (1982) propose, thought processes are not solely based on language and may be based on imagery. Accepting this, Ericsson and Simon (1980, 1984) affirm that individuals may not have access to all the processes involved, but they state that what individuals are saying has meaning and importance to them and is likely to be relevant when making decisions. To support this, they present empirical evidence to show that the kinds of thoughts verbalised are sufficient to account for the cognitive mechanisms when performing a task (Ericsson and Simon, 1993)

Despite these criticisms, there is evidence to suggest that the approach has validity. Using multivariate analysis, Bolton (1991; 1993) identified four underlying dimensions to individuals' responses. These dimensions represent the cognitive difficulties of comprehension, retrieval, judgment and response. The stability of these across different domains attests to the validity of the approach in identifying cognitive difficulties. Bickart and Felcher's (1994) findings add support for a four stage model of question answering for behaviour frequency questions. Evidence of validity also comes from studies that have used verbal protocols to identify defective questions and to modify questions. Bolton (1991; 1993) compared versions of a questionnaire and was able to test and identify where participants experienced difficulty. From this different versions were compared to improve the survey. Similarly, DeMaio *et al.* (1993) demonstrated that changes to questions as a result of cognitive interviewing improved the recall of dietary intake.

### **3.3.3 Limitations of verbal protocols**

One criticism of verbal protocols is that individuals may intentionally report their thoughts inaccurately because of presentational biases (e.g. to be seen in a positive light) (Cacioppo *et al.*, 1997). One way to minimise this is to reduce demand characteristics in the interview. In the present study it is unlikely that participants would need to distort the processes they used to report on patients' symptoms. However, in order to reduce self-presentational biases the instructions to proxies will stress that there were no right or wrong answers. Proxies will be

instructed to verbalise their thoughts even if they seemed muddled and not to try to sort their thoughts into any order.

Another limitation of verbal protocols is that they rely on the individual's ability to verbalise their thoughts. Those with poor verbal skills or cognitive limitations may have difficulty conceptualising and verbalising their thoughts. There is some evidence to suggest that older people are less able to verbalise their thinking than younger people (Jobe & Mingay, 1990). However, in this study the use of verbal probes was useful in older participants for identifying problems with questions (Jobe & Mingay, 1990). Age related problems were not specifically addressed in these investigations. However, difficulties with responses will be observed and noted.

The focus of cognitive interviews is on information processing; within palliative care research there is a need to think about emotional aspects of question answering. In an attempt to address this, reference to proxies' own feelings will be explicitly coded. This issue will also be directly addressed by research question 3, where the extent to which proxies' own emotions had a bearing on their reports is investigated.

### **3.4 Analysis of verbal protocols**

The analysis of verbal protocols varies according to the researcher's theoretical background and the research objectives. In most cases, investigators have performed content analysis as a means of interpreting the transcripts of cognitive interviews. They have also quantified the codings to reveal underlying difficulties that individuals have encountered when answering a question (Bolton, 1993; Bolton & Bronkhorst, 1996) or to reveal the strategies individuals adopt when reaching a decision (Blair & Burton, 1987; Bickart *et al.*, 1990). These codes are often based on a theoretical scheme that incorporates comprehension, retrieval, judgment, and response difficulties (Bolton 1993; Willis *et al.*, 1991; Sudman *et al.*, 1996). The coding scheme is used as a framework to deductively assign codes to the transcripts. Before discussing the analysis in detail, it is important to highlight the coding scheme used in verbal protocols. This is very important because it has implications for the type of analysis chosen.

#### **3.4.1 Coding verbal protocols**

Typically, the standard unit of analysis is the entire verbalisation (Willis *et al.* 1991). However, segments or words can be coded individually if required. To enable the identification of macroprocesses, the speech burst or segment was chosen as the unit of analysis for these investigations. Since there is no published research to date using verbal protocols in palliative care research, the analysis of the transcripts and development of the codes will be both inductive and deductive. This is very similar to ethnographic content analysis, where some code

categories are fixed, while others emerge through the study (Altheide, 1987). This is important because restricting the analysis to just fixed categories might exclude information that may be particularly relevant to palliative care research. Since this an entirely new investigation, wherein not all of the parameters are known in advance, it is important to maintain some flexibility within an otherwise structured coding scheme.

### 3.4.2 Coding scheme

The coding scheme is used to capture information about the processes involved in answering a question, such as how participants arrived at a response. However, content-based information can also be coded. Content-based information, as Bickart and Felcher (1996) highlight, can be valuable in understanding response processes as it can facilitate in the identification of information the participant uses to base their decision. This can be normative behaviour (what is deemed normal), trait knowledge (knowledge based on the person's characteristics/personality) and general knowledge. For example, proxies might centre their assessment of the patient's level of depression on the patient's usual character. The types of information described are more likely to occur when participants are asked to report about others (Blair *et al.*, 1991; Menon *et al.*, 1995). Therefore, content-based information was incorporated into the coding scheme in these investigations.

In addition to this, content-based information can also be used to identify response difficulties. For instance, certain words expressed by the participant, such as "*define*" and "*not familiar*" may be indicative of uncertainty or comprehension problems (Bolton, 1993). Coding response difficulties is important in this study as a means of identifying defective questions and for identifying the limitations of using proxies; so this was included in the coding scheme.

The coding scheme developed for these investigations was based on one developed by Sudman *et al.* (1996), which was used to assess behavioural and attitudinal questions with participants and their spouses. Categories represent comprehension, retrieval and judgment problems. It relies on the coder(s) to interpret the text and make inferences about the thought processes involved. This was chosen because it allows for the diagnosis of potential problems both with the questionnaire and with answering the questions. The coding scheme will be reviewed in more detail in Chapter 5 with the analysis of the verbal protocols.

### 3.4.3 Reliability of the coding

The analysis can be automated to improve reliability and to reduce the time and costs (Bolton & Bronkhorst, 1991). This takes the form of content analysis where the computer finds words or phrases predetermined and programmed by the researcher (Stone *et al.* 1966; Bolton &

Bronkhorst, 1991). However, for these investigations manual coding was chosen because inductive and deductive coding development was used. Furthermore, manual coding captures the richness of the verbal report that cannot be captured using automatic coding. In previous investigations the percentage of matches between coders have been calculated with average percentages across questions as high as .77 (Bickart & Felcher, 1994). This does not account for chance agreement between coders so the kappa statistic was chosen as a reliability estimate here. Kappa has been used to examine the level of agreement on behavioural frequencies, where estimates of .75 to .90 have been reported (Fowler & Cannell, 1996). The reliabilities of the coding will be reported in the relevant chapters following the analysis.

### **3.5 Content analysis**

Analysis of cognitive interviews follows the traditional content analysis where large volumes of information are reduced in to a smaller number of categories with similar meanings (Weber, 1985). Predominantly, the codes have been quantitatively analysed. Here the emphasis is on counts and frequencies where greater frequency is equated to greater significance. Statistical techniques such as factor analysis have also been applied to the codes to reveal underlying themes (Bolton, 1991; 1993). However, verbal protocols also lend themselves to a more qualitative approach. Content analysis can be used in this way as Weber, (1985) states *“the best content analytic studies utilize both qualitative and quantitative operations on text”* (p.10).

#### **3.5.1 Quantitative analysis of content**

Examining the frequency of words and phrases can be useful for discovering shared meanings. For example in surveys it is important that terms used are understood in a similar way to allow for interpretation of the findings (Willis *et al.*, 1990). Divergent interpretations may indicate that the question or terms are not suitable because in effect participants are addressing different questions. Consequently, this was used to investigate proxies' understanding of the pain, anxiety and depression frequency and severity descriptors. Content analysis is also useful for making comparisons between groups based on characteristics, for example codes can be compared between males and females. Research question 3 sets out to investigate whether proxies' own feelings have a bearing on their accounts. Here the groups can be subdivided based on their self-reported anxiety, depression and caregiver strain. A content analysis would enable a comparison of codes between those with high and low levels on these measures.

### 3.5.2 Qualitative content analysis

In contrast to traditional content analysis there is more flexibility for codes to be modified or new ones added during the analysis. In this analysis codes are developed from the coding scheme and inductively from the text. Furthermore, the analysis goes beyond quantification of the text to interpretation of the meaning behind what individuals are saying. Or as Morgan (1993) posits, from questions about *what* and *how many* during coding and counting, to *why* and *how* during qualitative content analysis. Content analysis was used in this way in these investigations because both approaches can contribute to understanding the text more fully.

From an epistemological standpoint content analysis comes from a realist perspective, where it is assumed that there is some fit between the outcome of the data analysis and reality. However, it is considered more of a method that resembles quantitative approaches when analysing frequencies. Whereas, when used qualitatively it leans toward more interpretative approaches such as grounded theory (Morgan, 1993). Although it should be noted that it has more in common with classic grounded theory (e.g. Glasser & Strauss, 1967; Glasser, 1978) where realism is reflected in participants' accounts and can be approximated from them (Chamberlain, 1999). Later versions of grounded theory (e.g. Charmaz, 1995; Strauss & Corbin, 1994), as Chamberlain (1999) identifies have moved away from realism into a constructivist epistemology. From a social constructivist perspective the true meaning of a phenomenon is thought not to exist. Instead, it is believed that there are many versions of reality. Thus, the qualitative inquiry is interested in how the phenomenon is constructed rather than a search for a reality (Potter & Wetherall, 1987; Parker, 1992).

### 3.5.3 Ethics of cognitive interviews.

This section does not deal with procedural matters such as ethical approval or support, as this is covered in the procedures section (Section 3.6.1). Instead it is specific to the methods employed and the ethics of using cognitive interviews with bereaved relatives. Morally there is a limit to the amount of verbal probing that can be used, which limits the use of the cognitive interview in the field of palliative care. However, the method is particularly well-suited to the research questions. As a researcher and a nurse practitioner my experience and judgment was used to guide when it was no longer appropriate to continue a particular line of questioning, even if this compromised the research aims. An added concern was the focus on particular aspects of the proxies' experiences based on the researcher's framework. Given the structured nature of the interviewing, this could have excluded a fuller understanding of the experiences from the proxies' point of view. Therefore proxies were not discouraged from expressing their views and relevant field notes were kept and used in the interpretation of events. Further still was the concern about asking the proxy to think about issues that they



may not want to think about or had not thought about before. For example, the extent of the patient's suffering with regard to pain, anxiety and depression. In line with ethical guidelines for research with human subjects proxies were reminded of their rights regarding participation in research studies (British Psychological Society, 2000). Alternatives to bereaved relatives were also considered, amongst these were the use of other groups such as proxies for older individuals. Despite the similarities in the patient-proxy literature, the specific circumstances and populations studied may not generalise to the retrospective approach for patients at the end-of-life. Hence, the use of bereaved relatives was deemed essential.

## **3.6 Design and procedures**

Since the research questions are addressed within one study, the design and procedures for all the investigations are detailed in this section.

### **3.6.1 Ethics of interviewing bereaved relatives**

Ethical concerns have been raised regarding the use of bereaved relatives in research and ethical guidelines proposed for conducting such research (Parkes, 1995; Robertson *et al.*, 1997). These concerns were taken into consideration when designing the present set of investigations since the participants are being asked to recall potentially distressing information at a very sensitive time. Before seeking ethical approval, a number of ethical safe guards were put into effect. This included clinical supervision and social work support from the hospice involved in the care of the patient and the proxy. In anticipation of potential distress caused by the interview, an ethical protocol was devised after negotiations with the hospice. A summary of the ethics protocol is contained in Appendix B.

Ethical approval was then sought and obtained from three Local Research Ethics Committees. With the agreement with the hospice's internal ethics group and the hospice social work department, it was decided that the social worker responsible for the area in which the patient resided, would screen out those individuals that were unsuitable to take part. In total, four individuals were excluded from the investigation; the reasons for exclusion included another bereavement, illness in the family or overwhelming distress. Although this does raise questions concerning sample bias, moral concerns for the welfare of distressed individuals took precedence.

An invitation letter was sent to the next of kin of those identified. This letter included information about the study and the rights of participants, along with a consent form (Appendix C). An opt-in approach was adopted, which meant that only those wishing to take part needed to reply to the letter. A telephone number was given if participants required further information or clarification. Participants who agreed to take part were then contacted by tele-

phone to arrange an interview. Participants were given the choice to be interviewed in their own homes or in a quiet place of their choosing. All of the participants chose to be interviewed in their homes. Prior to being interviewed, participants were reminded of their rights as set out in the British Psychological Society's Ethical Guidelines for research on human subjects (2000). Written and verbal consent was obtained to audiotape the interview. Participants were assured that the tapes would be destroyed after transcription.

### **3.6.2 Sample**

The sample included all the next of kin of patients who were known to one hospice in south-east London. The hospice covers a wide geographical area and takes in both urban and semi-rural locations. Participants were identified through a database held in the hospice's Social Work Department and included all the next of kin of patients who had died between February 2000 and April 2000. The inclusion criterion was: bereavement in the last three months; cognitively intact; contact with the patient in the last month and week before the patient's death. Participants also had to live within greater London, Surrey or Kent because of travel to and from interviews.

### **3.6.3 Measures and procedure**

Demographics and characteristics associated with the patient and proxy relationship were collected initially. These included details regarding the amount of contact they had with the patient, their relationship to the patient, the patient's cause of death, place of care and where they had died (Appendix D). Following this, participants were asked to complete three self-report instruments; Caregiver Strain Index (Robinson, 1983), State Anxiety Inventory (Spielberger *et al.*, 1983) and the Beck Depression Inventory (Beck *et al.*, 1996). These assessments were included as part of the investigation of research question 3 in which proxies' emotions are examined in relation to their reports of patient's pain, anxiety and depression.

The instruments were self-administered to 28 participants and orally to two because of visual impairment and illiteracy. The measures were presented in the same order to all participants and all were administered before the interview. This avoids possible carry-over effects from the interview that might affect the proxies' responses. This is particularly important for the anxiety instrument because it is context dependent.

#### ***Caregiver Strain Index***

The Caregiver Strain Index (CSI) developed by Robinson (1983) was used to evaluate the negative effect of caring for the patient (Appendix E). The instrument is short and easy to administer, with thirteen items assessing areas associated with the burden of caring. Summing all the positive responses from the dichotomous response options scores the scale. The maxi-

maximum possible score is 13. Scores of 7 or above indicate greater strain. Evidence for the instruments construct validity comes from three areas: caregivers' subjective perceptions of caring, ex-patient characteristics and the caregiver's emotional health. Robinson (1983) has shown that the instrument has high internal consistency with Cronbach's alpha of .86 reported. Proxies took less than 5 minutes to complete the instrument.

### ***State Anxiety Inventory©***

From the State-Trait Anxiety Inventory (Form STAI Y-1) (Spielberger *et al.*, 1983), only the State Anxiety Inventory (SAI) was used to measure the severity of proxy's anxiety. The SAI assesses how the participant feels at that particular moment and thus is assessing transitory context dependent anxiety. Thus, the test-retest reliability for SAI is understandably low (.16-.62). The instrument consists of 20 items scored on a four-point scale from *not at all*, *somewhat*, *moderately so* to *very much so*. Ten items assess the presence of anxiety, for instance, "*I feel frightened*" and "*I feel indecisive*". For these items, a rating of 4 indicates the presence of high levels of anxiety. The other 10 items indicate the absence of anxiety, for example, "*I feel satisfied*" and "*I feel steady*". These items are reversed prior to scoring. Overall, scores range from 20 to 80, with higher scores signifying greater anxiety.

There is good support for the reliability and validity of the instrument in clinical and non-clinical populations. Internal consistency rates of .93 attest to its reliability. Spielberger *et al.* (1983) has demonstrated both the discriminative and construct validity of the STAI. Norms are available for the instrument. The instrument took participants 6-10 minutes to complete.

### ***Beck Depression Inventory©***

The latest edition of the Beck Depression Inventory (BDI), the BDI-II was used to assess the presence and level of depressive symptoms over the last two weeks (Beck *et al.*, 1996). The instrument consists of 21 items that cover symptoms of depression, for example, mood, pessimism, suicidal thoughts and feelings of guilt. This latest version brings the instrument in line with the current depression criteria of the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM-IV, 1994). Each item is graded in intensity on a four-point scale with scores ranging from 0 to 3. The reliability and validity of earlier versions of the BDI have been widely reported in clinical and non-clinical populations (Beck *et al.*, 1988). However, the developers state that "*the BDI-II constitutes a substantial revision of the original BDI*" (Beck *et al.*, 1996, p.1). The psychometric properties of the BDI-II have been evaluated in a number of studies conducted by the developers. They report internal consistency rates of .92 in clinical and .93 in non-clinical populations have been reported, along with test-retest reliability of .93. This implies that the instrument has good reliability over time

(Beck *et al.*, 1996). The developers have also confirmed that the instrument has good convergent and discriminative validity. In this study the instrument took between 5-10 minutes to complete.

### ***Proxy Semi-structured Cognitive Interview***

The Proxy Semi-structured Cognitive Interview (PSCI) was developed by the investigator as a means of organising the gathering of information from proxies (Appendix F). It is a flexible design that allows the recall of as much information as possible from the proxy's perspective. As the name suggests, the PSCI contains components that are both structured and unstructured. The structured component asks participants to respond to questions regarding patient's pain, anxiety and depression in the last week and month before death. These questions are drawn from VOICES, although it should be noted that there are differences. These are detailed in Appendix G. The unstructured components of the PSCI leave cognitive questions deliberately open-ended so that participants have the opportunity to express themselves fully. The thrust of the cognitive questions is to ascertain how they arrived at their answer, by encouraging participants to introspect and verbalise their thoughts (discussed in Section 3.3.1). It is hoped that some insight will be gained into how individuals represent memories for events and how they arrive at their responses.

Interviews took between 35-75 minutes. Intervention by the interviewer was kept to a minimum in order to avoid "*cueing*" the participant. Appendix H contains information on the prompts used. The interviews were taped recorded and transcribed verbatim. Following each interview, the interviewer noted their own thoughts and feelings. This was a way of distinguishing between those of the participant and those of the interviewer.

### **3.6.4 Pilot study**

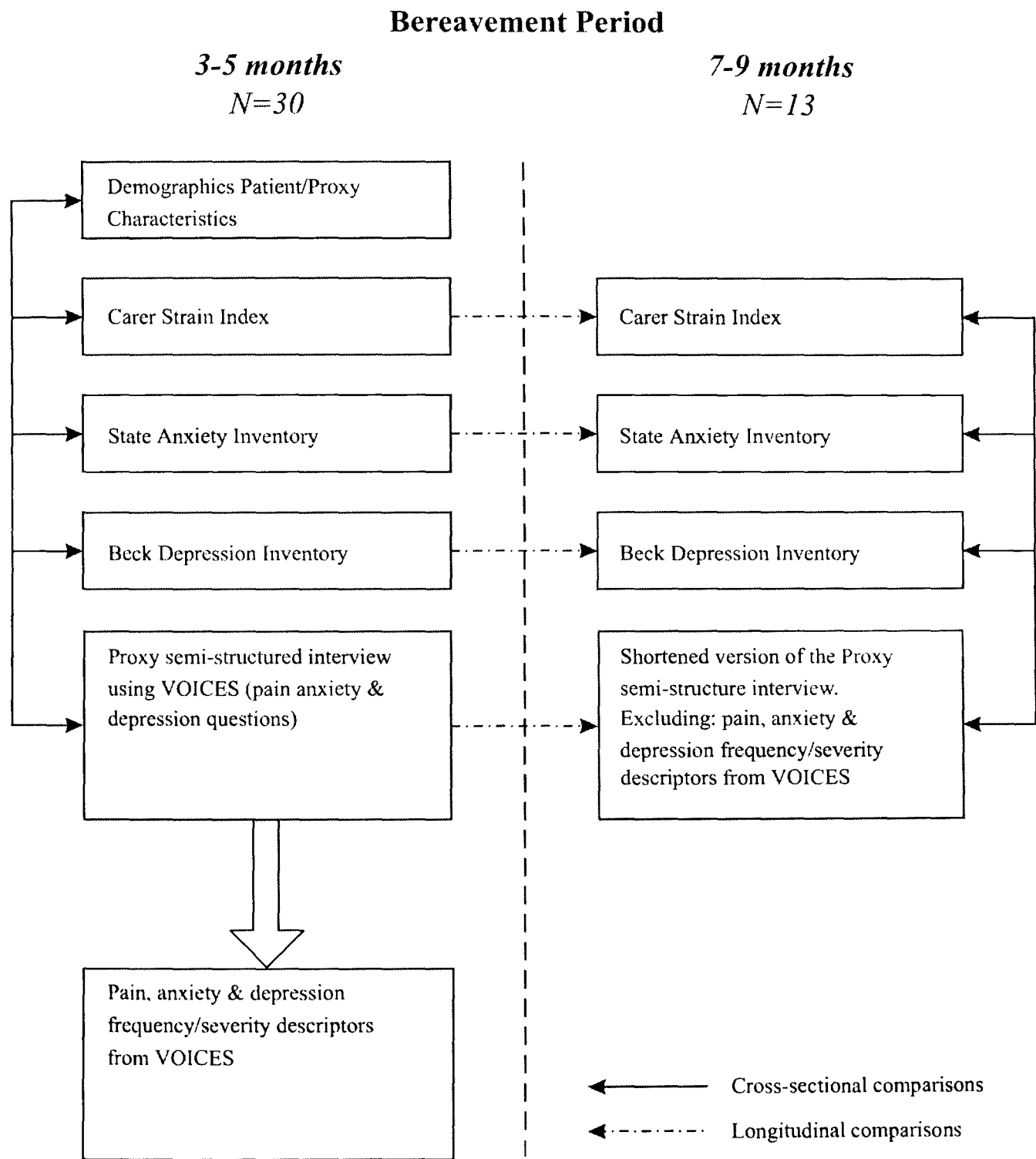
A pilot study was conducted with four participants. The purpose of the exercise was to check the comprehensibility and appropriateness of the questions, the procedure and the length of the interview. Few modifications were needed. Appendix I contains a list of the modifications and the reasons why they were made.

### **3.6.5 Follow up study**

Initially, participants were interviewed 3-5 months (mean 3.9 months) following the patient's death (Time 1). In order to investigate changes over time some participants were also interviewed a second time (Time 2). Interviews were conducted 3-5 months (mean 4.6 months) after Time 1, approximately 7-9 months (mean 8.5 months) after the patient's death. To reduce any carry over effects from the interview at Time 1, participants were not aware that they would be invited for a second interview. As for the first interview, the relevant Local

Research Ethics Committee gave ethical approval. An invitation letter was sent to those who participated at Time 1, this is contained in Appendix J. To keep the interview to a reasonable length for participants, the PSCI was shortened by taking out questions that related to the pain, anxiety and depression frequency and severity descriptors. A shortened version of the PSCI is in Appendix K. Participants also completed the CSI, SAI, and BDI. There was also no need to gather demographic information as this was already obtained at Time 1. Otherwise, the procedure and measures were the same as for the interviews at Time 1. Interviews took between 20-40 minutes. Figure 3.1 gives an overview of the entire research design.

**Figure 3.1** Summary of research design.



# CHAPTER 4: Frequency and severity descriptors

## 4.1 Introduction

A major objective of palliative care is to address quality of life issues affecting terminally ill patients. Control of pain and other symptoms experienced by the dying patient, including psychological, social and spiritual aspects are paramount (WHO, 1990). The management of acute and/or chronic symptoms prevalent among patients at the end of life is very challenging (Conill *et al.*, 1997; Curtis *et al.*, 1991; Donnelly & Walsh, 1995). This is particularly true for patients with advanced cancer, where as many as 70% experience some level of pain (Brescia *et al.*, 1992). Based on studies of hospice patients suffering from terminal cancer, up to half experience moderate to severe pain (Vainio & Auvinen, 1996). Of the psychological symptoms, severe anxiety and depression are two of the most common to occur in the advanced stages of the disease (Derogatis *et al.*, 1983; Roth & Breibart, 1996). Prevalence rates for depression as high as 58% have been reported in terminally ill patients (Hinton, 1972), although others have reported lower levels of a quarter of all patients (Massie, 1989). There is no doubt that for some patients' feelings of fear are typical and can reach significant intensity (McDonnell, 1989; Hardman *et al.*, 1989). Levels of anxiety have been found to vary between 21%, and as high as 43% in terminally ill cancer patients (Coyle *et al.*, 1990; Addington-Hall *et al.*, 1992).

Appropriate management of symptoms in patients at the end of life has moral implications because it can ease suffering and limit their effects on quality of life. This is true not only for the patient, but for their family as well. It has been shown that those closest to the patient, significant others, share the patients' experiences (Ferrell *et al.*, 1991). Finding and applying the most appropriate measures for managing the symptoms of terminally ill patients can be very difficult, because many symptoms are interrelated and affect different patients in a variety of ways. For example, it is well recognised that there is an interrelationship between anxiety and depression (Clark, 1989; Gotlib & Cane, 1989). Moreover, these psychological indicators can often be related to levels of pain. There is substantial evidence to show that depression and chronic pain are closely interlinked (Rudy *et al.*, 1988). This relationship is recognised in some assessments of pain where mood is included in the measures. A good example is the McGill Pain Questionnaire (MPQ) where the affective component of MPQ has been shown to correlate with depression scores in a cancer population (Kremer & Atkinson, 1981). Anxiety and depression may also be a consequence of uncontrolled pain (Zimmerman *et al.*, 1996). Then again, pain may be an expression of psychological problem. Depression can manifest as irretractable pain, while anxiety can manifest as nausea or dyspnea (Barraclough, 1997). Thus any comprehensive evaluation of palliative care should include an

assessment of pain, anxiety and depression. However, as this thesis has identified agreement between patients and their significant other proxies is poorest for these symptoms.

One of the aims of this Ph.D. is to investigate the frequency and severity descriptors used in the VOICES retrospective survey (Addington-Hall *et al.*, 1998). It is important to establish what proxies understand regarding the terms used in interviews and questionnaires. The instrument must also provide a framework in which participants can accurately report patients' experiences and obviate situations where divergent interpretations or ambiguities can arise. Asking participants what phrases or terminology means can help in identifying difficulties with terms and can also identify common alternative wording used by participants (DeMaio & Rothgeb, 1996). Being aware of these potential problems and reducing them is an important first step to improving reliability and validity of the information provided by proxies. Before detailing the research in this study it is important to locate the present study in the wider context of pain, anxiety and depression assessment at the end of life.

#### 4.1.1 Pain

Pain is more than an aversive sensation. It is a complex, multidimensional phenomenon incorporating psychological, physical and sociocultural aspects (Ahles *et al.*, 1983; Ahles & Martin, 1992; Melzack, 1983). It is commonly experienced at the end of life for a variety of reasons. These include physical aspects of the disease such as cerebral metastases, or as a consequence of the disease on functioning (Breitbart & Holland, 1988; Breitbart *et al.*, 1995). Assessment of pain, such as its frequency and intensity, is central to pain control. Numerous instruments have been developed to assess pain. Those used for patients at the end of life include the McGill Pain Questionnaire (Melzack, 1975; Melzack, 1987), Memorial Pain Assessment Card (Fishman, 1987), The Edmonton Symptom Assessment System (Bruera, 1991), and The Wisconsin Brief Pain Questionnaire (Daut, *et al.*, 1983).

Assessment of pain can be multidimensional or unidimensional. A multidimensional pain instrument widely used in palliative care is the McGill Pain Questionnaire (Melzack, 1975; Melzack, 1987). The MPQ contains seventy-eight pain descriptors that are categorised into twenty subcategories. These subcategories reflect different dimensions of pain; sensory, affective and evaluative. The sensory dimension includes words that reflect the sensory nature of the pain, such as "*dullness*" or "*temporal*". Emotions, such as anxiety, are encapsulated in the affective component. The third dimension includes words that describe the intensity of the pain such as "*excruciating*". Unidimensional pain instruments tend to focus on one aspect of the pain, for example, pain intensity or frequency. For example, The Edmonton Symptom Assessment System (Bruera, 1991) only evaluates pain severity. Whatever type of assessment is used, the advantages and disadvantages of each instrument have to be balanced against the

ability of the individual to complete the measure and the goals of the research. In some instances, multidimensional scales may not be appropriate as they are time consuming and require good verbal skills (Briggs & Closs, 1999).

In retrospective studies of terminally ill patients, the VOICES survey is unidimensional. Since proxies may not have the necessary knowledge to report on the specifics of the pain experience, using unidimensional instruments is more common. VOICES assesses pain frequency and intensity using a five point verbal rating scale (VRS). Frequency is rated from “*most of the time*” to “*never*” and intensity is rated from “*severe*” to “*none*”. Other types of scales include visual analogue scales (VAS) and numerical scales. Typically, VAS indicates pain intensity by marking along a line 100 mm or 10 cm from no pain to the worst possible pain (Huskisson, 1974). Numerical scales are rated from 0-10 or from 0-100 along a continuum from “*no pain*” to “*most severe pain*” (Karoly & Jensen, 1987). A concern with VRS is their psychometric properties, as the intervals between each category cannot be assumed to be equal. Instead they are considered ranks (Heft & Parker, 1984; Littman, *et al.*, 1985). VRS also require an interpretation of the categories to assess whether the experience can be represented in the scale. Therefore, language is an essential feature of VRS and an understanding of how these descriptors are interpreted is paramount.

#### **4.1.2 Anxiety and depression**

Psychological symptoms such as anxiety and depression can occur because of physical aspects connected with the disease such as metabolic imbalances or medications (Breitbart & Holland, 1988; Breitbart *et al.*, 1995). They can also occur as a result of loss associated with the illness and its course. Fears about the future, uncontrollable symptoms and a decrease in functional abilities are recognised as important features in the development of psychological problems (Silverfarb *et al.* 1980; Breitbart & Jacobsen, 1996). The Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV, 1994) classifies a number of symptoms that characterise anxiety, and mood disorders such as depression. Features common to both conditions are sleep disturbance, fatigue, difficulty concentrating and significant distress or impairment of social, occupational or other areas of functioning. This overlap highlights the difficulty proxies may encounter when deciding whether the patient was anxious or depressed. Moreover, the symptoms of fatigue, weight loss and poor concentration are common in patients with cancer (Chochinov, 2000). Thus, identifying depression and anxiety may be difficult.

Differentiating features of anxiety and depression, as the DSM-IV (1994) illustrates, are based on few observable aspects. The individual’s thoughts and feelings best characterise anxiety from depression. Suicidal ideation, worthlessness and sadness are specific to depres-



sion, whereas worry, apprehension and expectation are particular to anxiety. This can make assessment of the symptoms by others difficult and may account for under treatment (Craig, 1989). Further to this, these reactions may be considered a normal part of adapting to a life-threatening illness.

Assessment of anxiety and depression in palliative care usually relies on patients' self-report using measures such as the Beck Depression Inventory (Beck *et al.*, 1961), General Health Questionnaire (Goldberg, 1972), Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983) and the Profile of Moods (Cella *et al.*, 1987). Similar to pain assessment, various scales have been used to assess psychological symptoms. The instruments themselves vary; some assess anxiety and depression using symptoms set out in the DSM-IV to evaluate the level of anxiety or depression. A good example is the Beck Depression Inventory (Beck *et al.*, 1961). The instrument consists of 21 items that covers symptoms of depression, namely mood, pessimism, suicidal thoughts and feelings of guilt. This assessment contrasts to instruments that assess multiple symptoms where mood is assessed on a single item scale such as intensity. For instance, the Support Team Assessment Schedule (Higginson & McCarthy, 1992) where patient anxiety is rated on a four-point scale. These rely on the patients' interpretation of the term anxiety.

Retrospective assessment instruments like VOICES ask the proxy to assess the presence, intensity and severity of anxiety and depression. For example VOICES asks "*How severe was their level of anxiety during the last week of life?*" This is rated on a five-point scale from "*severe anxiety*" to "*none*", as is depression. However, little is known about how proxies interpret these descriptors and the language used. Another approach has been to use the POMS anxiety and depression subscales (Lynn *et al.*, 1997; Somogyi-Zauld, 2000). POMS ask the participant to rate statements of adjectives pertaining to mood state, such as "*unhappy*", "*discouraged*" on a five-point scale from 0 (not at all) to 4 (extremely). Here knowledge of the symptoms of anxiety and depression is not required although the instrument is more time consuming to complete. One criticism of this type of measure is its ability to capture the range of symptoms that accompany a psychological state (Snaith, 1981). This might be a particular concern when scales are shortened. There is evidence to suggest that single item scales such as VAS are comparable to POMS for detecting depression (Kilgore, 1999). In retrospective surveys of palliative care proxies are required to report on various aspects of the care; sources of informal and formal care; service utilisation and evaluation; information and communication with health care professionals; patients symptoms and restrictions, and the proxy's experience of bereavement and bereavement care (VOICES, Addington-Hall *et al.*, 1998). In order to keep this to a manageable length single-item scales are preferred. However, this should not compromise the accuracy of the information from proxies. Conse-

quently, it is necessary to investigate proxies understanding of the terms used in single item scales such as VOICES. As Leff (1978) identifies, words like depression may mean different things to different people.

## 4.2 Aims

The aim of this investigation is to examine the pain, anxiety and depression frequency and severity descriptors used in the VOICES retrospective palliative care survey (Addington-Hall *et al.*, 1998). This is a way of identifying ambiguous quantifiers in the questionnaire and as a means of improving the validity of VOICES. This is important because “*symptom assessment and measurement is dependant on the clarity of meanings attached to symptom descriptors*” (Ingham & Portney, 1999, p.203).

### *Research question one*

What do proxies understand by the pain, anxiety and depression frequency and severity descriptors used in retrospective questionnaires that are used to assess palliative care?

## 4.3 Methods

The methods and procedures are described in Chapter 3. This section will detail those questions specific to these investigations. As part of the proxy semi-structured cognitive interview (PSCI), proxies were asked a number of open-ended questions designed to explore their understanding of the pain, anxiety and depression frequency and severity descriptors. Each of the scales was divided into five levels. These are *severe, significant, moderate, mild, none* for the intensity scale and *most of the time, often, sometimes, rarely, never*, for the frequency scale<sup>2</sup>. As an example, participants were asked, “*In your own words, what do you think to experience anxiety often means?*” and “*In your own words what comes to mind when you think of severe anxiety?*” In addition to this, participants’ rated descriptors on a numerical intensity scale from 0 (*none*) to 10 (*severe*). This type of magnitude estimation allows the researcher to examine the calibration of the scale. That is, do participants quantify the symptom in a similar way? What are the intervals between the descriptors on the scale? The questions analysed in this section are contained in Appendix L (questions 2c-2e, 3c-3i, 12c-12e, 13c-13i, 20c-20e, 21c-21i).

## 4.4 Analysis and findings

During a three-month period from May to July 2000, 121 cases were identified from records at the hospice. Of these, four were thought to be too distressed by their social workers to be

---

2. These descriptors are from an earlier version of VOICES. Later versions have changed the intensity descriptors significant and severe to severe and very severe.

contacted to take part. A further twenty-four were not contacted as they did not live in the area or had moved out of the area. Seven had no relative or contact addresses for a relative. This left eighty-six potential participants. From this, thirty-eight responded to the letter. Six replied stating that they did not wish to take part as their relative had been well cared for and did not have any pain, anxiety or depression. Two replied and consented to take part in the study, but could not be contacted for interview. Thirty participants were interviewed, thirty five percent of those contacted. Time from the patient’s death to the proxy interview was 3.9 months (SD .77) (range 2.43-5.17 months). Information on those who participated in the study and those who did not were examined for differences. Limited information was available for individuals who decided not to take part. No differences were found for patient age ( $t=-.097, p=.92$ ), patient sex ( $\chi^2=.40, p=.53$ ), proxy sex ( $\chi^2=.35, p=.55$ ), relationship between the patient and proxy ( $\chi^2=.49, p=.48$ ) and place of death ( $\chi^2=3.85, p=.15$ ) (full analysis contained in Appendix M).

### 4.4.1 Descriptive statistics

Descriptive statistics of patient and proxy characteristics were produced using SPSS/Win 10.1 and summarised in Tables 4.1 and 4.2. There were more female (19/30) than male (11/30) proxies in the sample interviewed. In contrast, there were equal numbers of female and male patients. The average age of the proxies was 59 years, younger than the patient’s average age of 71 years. There was a wide age range in both due to one patient/proxy dyad being a lot younger, 25 and 27 years respectively. Most of the patients and proxies were of caucasian western European ethnicity. They originated from diverse socioeconomic backgrounds covering all six social classifications (National Statistics Socio-economic Classification, 2000). All the proxies were close family members. Sixty percent (18/30) were partners. The amount of contact between the patient and proxy in the last week of life was high due to the fact that the majority lived with the patient (24/30) and cared for them during their last week (21/30). Despite the finding that half (15/30) the patients died in the hospice, many (6/14) of these spent less than 24 hours in the hospice prior to their death, with a further four patients in the hospice less than 72 hours. Prior to admission to the hospice they had been cared for at home. The predominant cause of death was cancer, with only one non-cancer death from coronary heart failure. Table 4.1 details the cancer types.

**Table 4.1: Patient demographics and characteristics.**

| Category | Criteria           | N | (%) |
|----------|--------------------|---|-----|
| Age      | Range 25-90 years  |   |     |
|          | Mean 71 (S.D.13.4) |   |     |

**Table 4.1: Patient demographics and characteristics.**

| Category                                      | Criteria                | <i>N</i> | (%)  |
|---|-------------------------|----------|------|
| <i>Gender</i>                                 | Males                   | 15       | (50) |
|   | Females                 | 15       | (50) |
| <i>Ethnic group</i>                           | European                | 27       | (90) |
|   | Hispanic                | 2        | (6)  |
|   | Afro-Caribbean          | 1        | (3)  |
| <i>Socio-economic status</i>                  | Class 1                 | 4        | (13) |
|   | Class II                | 3        | (10) |
|   | Class III               | 6        | (20) |
|   | Class IV                | 10       | (33) |
|   | Class V                 | 3        | (10) |
|   | Class VI                | 4        | (13) |
| <i>Cause of death</i>                         | Lung                    | 12       | (40) |
|   | Colorectal              | 4        | (13) |
|   | Upper gastro-intestinal | 4        | (13) |
|   | Breast                  | 3        | (10) |
|   | Prostrate               | 3        | (10) |
|   | Liver                   | 1        | (3)  |
|   | Brain                   | 1        | (3)  |
|   | Spine                   | 1        | (3)  |
|   | Coronary heart failure  | 1        | (3)  |
| <i>Place of care in the last week of life</i> | Home/hospice            | 9        | (30) |
|   | Home                    | 8        | (27) |
|   | Hospice                 | 4        | (13) |
|   | Hospital                | 4        | (13) |
|   | Hospital/home           | 4        | (13) |
|   | Nursing home            | 1        | (3)  |
| <i>Place of death</i>                         | Hospice                 | 15       | (50) |
|   | Home                    | 9        | (30) |
|   | Hospital                | 5        | (17) |
|   | Nursing home            | 1        | (3)  |

**Table 4.2: Proxy demographics and characteristics.**

|   |                           | N  | (%)    |
|---|---------------------------|----|--------|
| <i>Age</i>  | Range 27-88 years         |    |        |
|   | Mean 59 (S.D.13.6)        |    |        |
| <i>Gender</i>   | Males                     | 11 | (37)   |
|   | Females                   | 19 | (63)   |
| <i>Ethnic group</i>   | European                  | 28 | (93)   |
|   | Afro-Caribbean            | 1  | (3)    |
|   | Hispanic                  | 1  | (3)    |
| <i>Socio-economic status</i>  | Class I                   | 4  | (13)   |
|   | Class II                  | 5  | (16.5) |
|   | Class III                 | 8  | (27)   |
|   | Class IV                  | 5  | (16.5) |
|   | Class V                   | 5  | (16.5) |
|   | Class VI                  | 3  | (10)   |
| <i>Relationship to patient</i>  | Partner                   | 18 | (60)   |
|   | Son/daughter              | 10 | (33)   |
|   | Ex-partner                | 2  | (7)    |
| <i>Residence</i>  | Lived with patient        | 24 | (80)   |
|   | Did not live with patient | 6  | (20)   |
| <i>Amount of contact between the patient and proxy in the last week of life</i> | 24 hours per day          | 22 | (73)   |
|   | 12-24 hours per day       | 2  | (7)    |
|   | 6-12 hours per day        | 1  | (3)    |
|   | 3-6 hours per day         | 2  | (7)    |
|   | <3 hours per day          | 3  | (10)   |

**4.5 Numerical data analysis of the severity and frequency descriptors**

Proxies were asked to give a numerical rating from 0 (*none*) to 10 (*severe pain*) for the VOICES severity adjectives (*mild, moderate, significant*). This was to examine whether participants quantified the adjectives in a similar way.

Table 4.3 shows some consistency in the mean rating across the symptoms, with participants rating “*mild*” as 2, “*moderate*” as 5 and “*significant*” between 7-8. The quantities (0-10) assigned to the descriptors showed that moderate levels were approximated to the mid range of the scale. However, the standard deviations and range reveal that there was wide variability in the ratings particularly for *significant* and *moderate* pain and for *significant* depression. Furthermore, there was not always an equal interval between the points on the scale. For

example, on a 0-10 scale the average ratings for anxiety are not even. This is also evident in the overlap of the range of scores for all three symptoms.

**Table 4.3: Numerical ratings of severity adjectives for symptom severity in VOICES.**

| Symptom    | Pain |     |      | Anxiety |     |      | Depression |     |      |
|------------|------|-----|------|---------|-----|------|------------|-----|------|
| Quantifier | Sig. | Mod | Mild | Sig.    | Mod | Mild | Sig.       | Mod | Mild |
| Mean       | 7.4  | 4.9 | 2.3  | 7.8     | 5.0 | 1.9  | 7.6        | 5.1 | 2.3  |
| S.D.       | 1.2  | 1.4 | .89  | .79     | .87 | .61  | 1.0        | .92 | .81  |
| Range      | 4-10 | 2-8 | 1-4  | 7-9     | 3-7 | 1-4  | 4-9        | 4-8 | 1-4  |

## 4.6 Qualitative analysis of the severity and frequency descriptors

Further to the numerical analysis of severity adjectives, participants were ask to describe what they understood by the severity and frequency adjectives in VOICES. A content analysis on these short answer responses would enable comparison of similarities and differences in understanding. This was also a way of identifying ambiguous quantifiers in the questionnaire. All the texts were transcribed verbatim and entered into winMAX 98Pro, a qualitative computer analysis program.

An inductive content analysis was used to identify themes that emerged from the data. This involved coding the data. Codes, as Miles and Huberman (1994) describe, are “*tags or labels for assigning units of meaning to the descriptive or inferential information*” (p.56). Codes were attached to the response, segments or words. The codes were grouped into categories representing particular themes. In order to evaluate the comprehensiveness and appropriateness of the analysis, the codes and themes were discussed with another researcher. Furthermore, the data was independently coded to examine the reliability of the analysis. Cohen’s kappa was used to calculate the agreement between coders.

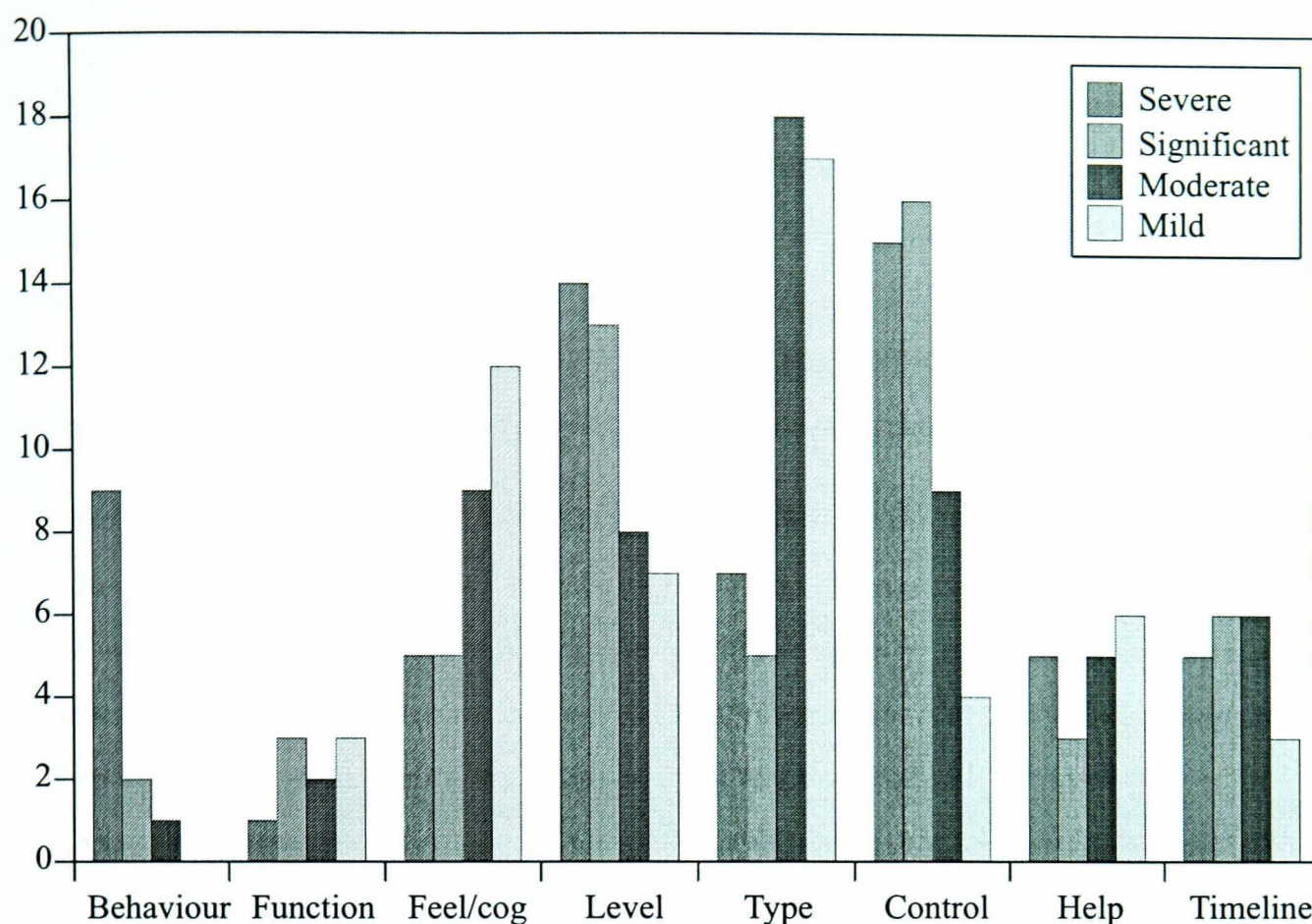
### 4.6.1 Findings from the pain severity descriptors

Analysis of the pain descriptors (severe, significant, moderate and mild) revealed 224 codes. These codes formed eight categories. Figure 4.1 details the frequency of each of the categories. The reliability of the codes was high with the strength of agreement ranging from  $\kappa = .94$  ( $p = .001$ ) to  $\kappa = .98$  ( $p = .001$ ).

*Behaviour* incorporated overt expressions of pain such as shouting and moaning. These were more likely at higher levels of intensity, as Figure 4.1 illustrates. This category included descriptions such as “*calling out in pain*” (ID 26) or “*in a really bad way moaning with it*” (ID 12). At moderate and mild levels behaviour was seldom described. Here the experience was less overt and experienced more internally as awareness of the experience. Although a person’s functioning can be described as an overt expression this was separated in this study



**Figure 4.1** Category frequencies for the pain intensity descriptors.



as functioning described more than a change in functioning. Here it was used to indicate lower levels of pain, as one proxy described mild pain, “*Not that painful, it doesn’t stop you carrying on with things*” (ID 6).

**Feelings/cognitions** included affective indicators of pain. This category was used at all levels of pain intensity. “*Distressing*” or “*disturbing*” was used to describe severe and significant pain ( $N=6$ ). An awareness of the pain across the levels was evident, from “*you definitely know it’s there*” (ID 15) at severe and significant levels of pain to “*moderate pain would be there you know it would be there*” (ID 3) at moderate levels. While at mild levels there was less awareness of the pain “*you forget you’ve got it after a while*” (ID 18).

**Tolerance** was the category most frequently referred to, especially at high levels of pain intensity (Figure 4.1). Tolerance referred to control or lack of control over the pain. This was subdivided into control which included coping with the pain, and help, the need for help with the pain or help received. Lack of control over the pain was evident at severe and significant pain levels with descriptions such as “*overwhelming*” and “*unbearable*” being used by a number of participants ( $N=8$ ). At the other end of the pain scale, control over the pain was to the extent that the pain was tolerable or bearable without the need for intervention. In contrast, help to control the pain was stressed as a need for intense levels of pain as many stressed that the pain would be difficult to cope with for any period of time ( $N=8$ ). Most interventions

mentioned were pharmacological, especially at intense pain levels. Whereas, a change in position or distraction from the pain were sufficient for mild pain.

**Intensity level** were markers or references used by participants to describe the level of the pain. Words such as “*excruciating*” and “*intense*” characterised the strength of severe pain ( $N=13$ ). Participants had more difficulty characterising significant pain. Here severe pain was used as a marker, with participants often ( $N=10$ ) stating that it was less intense than severe. For example, “*not so intense*” (ID 1) or “*it's similar to severe but not quite as bad*” (ID 12). At mild and moderate levels of pain participants made less reference to the level. Instead, the level of pain was more often characterised as an intensity type. These included descriptions of the pain such as a “*bruising*” or “*aching*” ( $N=25$ ). Several participants stated that at mild levels it was not really pain ( $N=9$ ). They characterised it as “*feeling uncomfortable*” (ID 25). In contrast, adjectives such as “*stabbing*” or “*sharp*” were used to characterise severe and intense pain ( $N=7$ ).

**Timeline** included the frequency of the pain like the chronicity, persistence or episodic nature of the pain. The category highlighted the overlap between symptom severity and frequency. At severe and significant levels pain was referred to as “*ongoing*” or “*constant*” ( $N=7$ ). Whereas, at lower levels of intensity it was often described as periodic or infrequent ( $N=18$ ).

#### 4.6.2 Findings from the anxiety severity descriptors

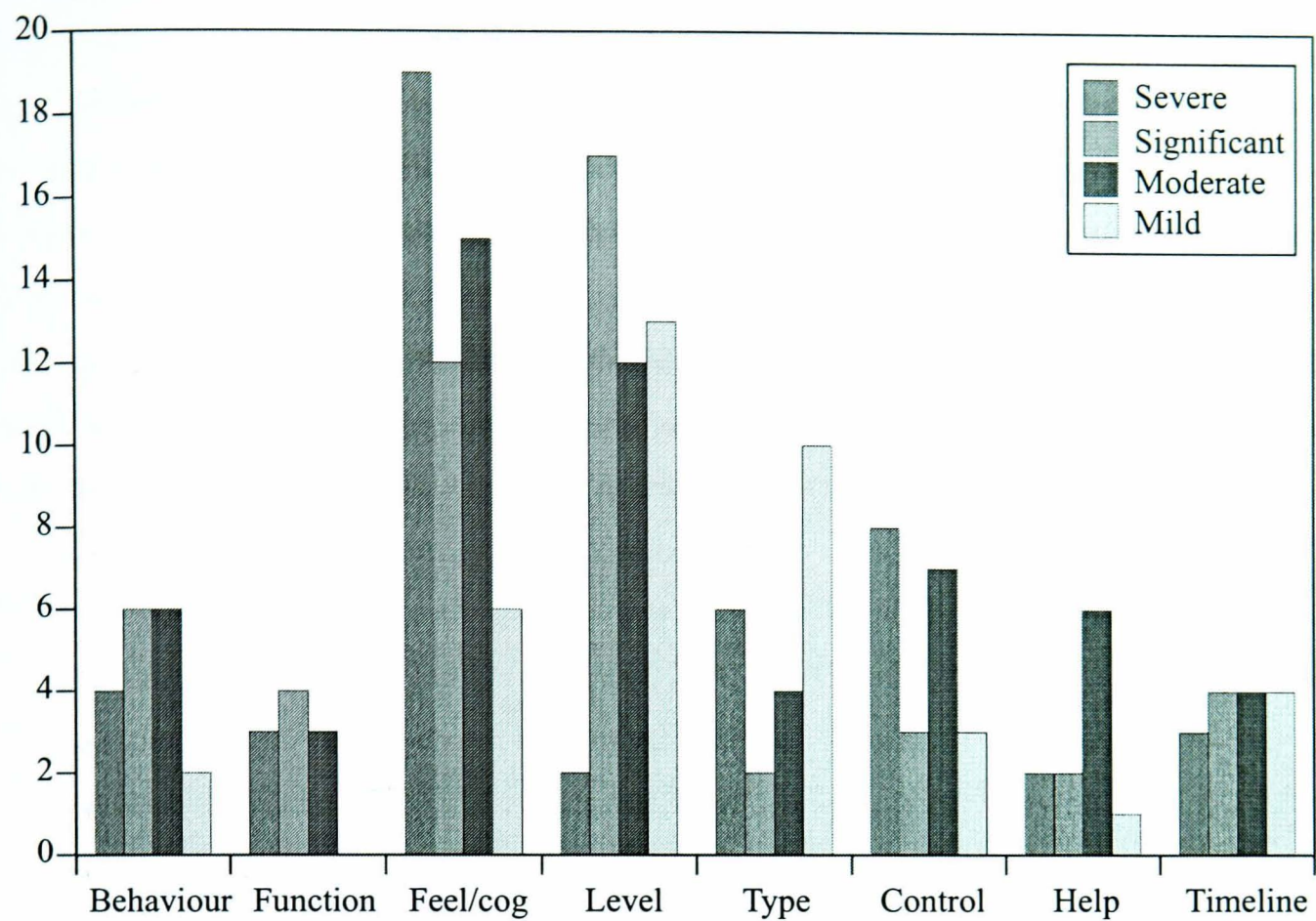
One hundred and ninety nine codes were identified in the anxiety descriptors (severe, significant, moderate and mild). These codes formed eight categories similar to the pain codes. Figure 4.2 contains the frequency of the categories. Examining the reliability of the codes, agreement was found to be near perfect ranging from  $\kappa=.82$  ( $p=.001$ ) to  $\kappa=.88$  ( $p=.001$ ).

Given the nature of anxiety references to thoughts and feelings in the category **feelings/cognitions** predominated in this analysis, as Figure 4.2 illustrates. Fears and overwhelming distress were most often mentioned at severe and significant anxiety levels with little abatement from these thoughts, “*you just can't focus on anything else*” (ID 15). At moderate levels the anxiety was still considered a problem but the focus had moved from the extreme fears associated with severe anxiety to a level where anxiety was seen to “*play on your mind*” (ID 24) and “*still very troubling*” (ID 10). While at mild levels many considered that the symptom was not a problem ( $N=10$ ).

Less reference was made to **behaviour** and **functioning** at all levels of intensity. Behavioural manifestations varied including “*agitation*”, “*panic attacks*” and as one participant stated “*generally low physical well-being*” (ID 21). These descriptions were evident at all levels. Codes relating to functioning were found in only ten instances. There was no men-



**Figure 4.2** Category frequencies for the anxiety intensity descriptors.



tion of functioning with mild anxiety. At moderate levels of anxiety participants varied, “let’s say worry that doesn’t stop you getting on with your life” (ID 18) while another thought of it “affecting your everyday life” (ID 29). At significant levels it was seen to impact on life. Whereas, severe anxiety was seen to greatly impact on functioning to the level of “incapacitating” (ID 25) and “almost disabling” (ID 30).

**Intensity level** was an important indicator of anxiety severity for significant, moderate and mild anxiety. Here severe anxiety was used as a anchor to reference the level of anxiety, “less than severe but still upsetting” (ID 1), “almost as bad as severe” (ID 21) and “in a similar league to with severe anxiety but to a lesser extent” (ID 13). **Intensity type** was less often used to characterise anxiety. Adjectives such as “jumpy” and “panicky” were used. At mild levels many participants used types to specify that the symptom was worry rather than anxiety (N=10), as one participant said, “Just minor worries that most of us deal with. Nothing like what I’d call anxiety” (ID 21)

In contrast to pain, the category **tolerance** did not feature in many of the participants’ replies. **Control**, that referred to control or lack of control over anxiety, at severe anxiety was viewed as insurmountable. Here “You’d have difficulty coping with it” (ID 14) and “at the limit of what you could take” (ID 18) characterised the intensity. Significant anxiety was viewed as still overwhelming but more bearable than severe anxiety. Whereas, moderate anxiety was a tolerable level, “more manageable” (ID 17). At mild levels participants felt that



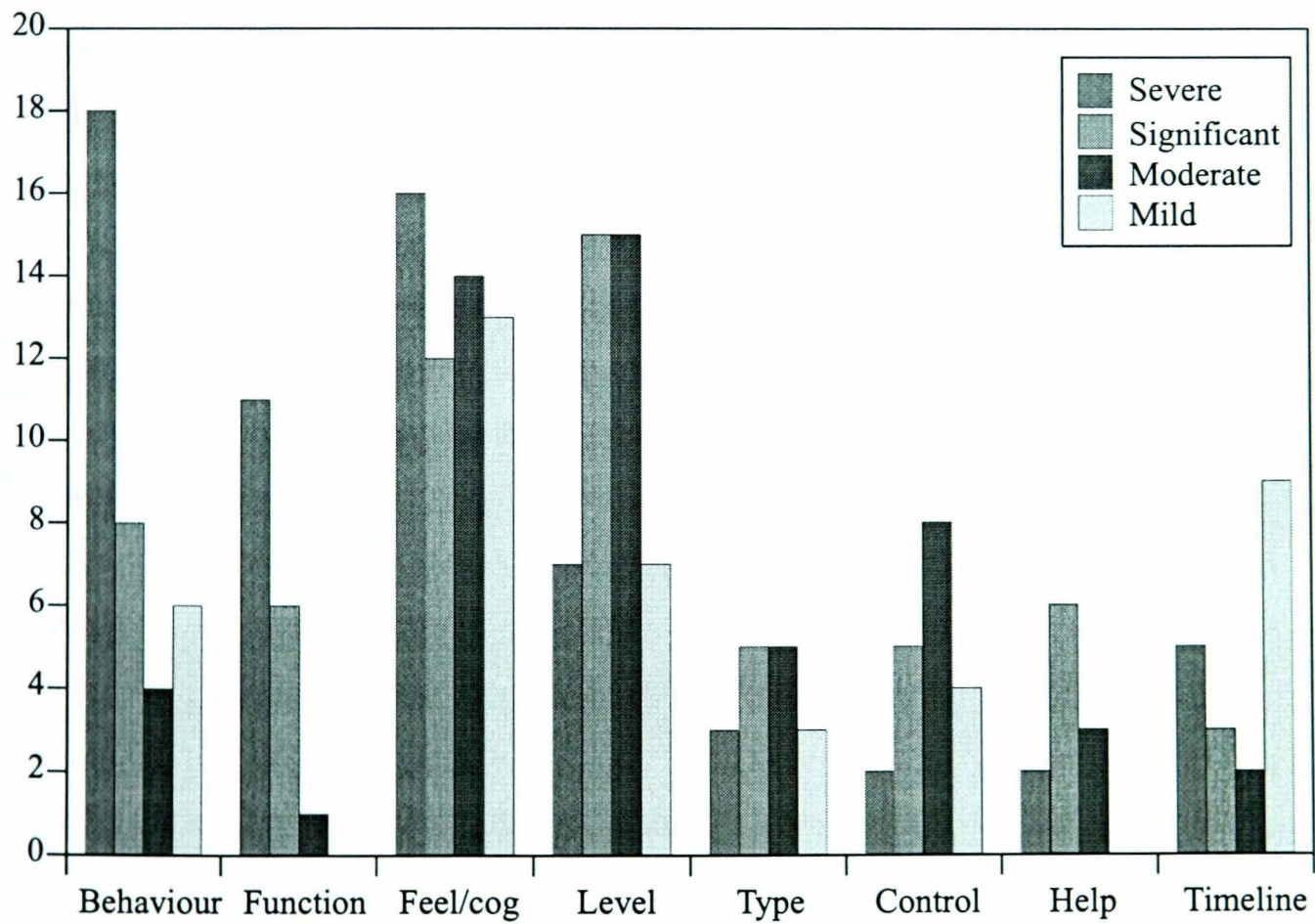
there was control over the anxiety ( $N=8$ ). Help, the need for help or help received tended to focus on the need for help rather than an intervention to alleviate anxiety. This might reflect a lack of knowledge of interventions.

**Timeline**, frequency of the anxiety overlapped with intensity. Severe anxiety was viewed as a persistent symptom “*constantly affecting you*” (ID 17). The timeliness of the anxiety varied at significant levels from having the symptom constantly or almost constantly to periodic episodes of intense anxiety. At moderate and mild levels the symptom was described as variable from times of feeling better to episodes of anxiety. A difference between moderate and mild anxiety was that several participants commented that mild anxiety was a reaction to an event, for example, an appointment or a letter ( $N=7$ ).

### 4.6.3 Findings from the depression severity descriptors

From an analysis of the depression descriptors (*severe, significant, moderate and mild*), 218 codes were revealed. The codes formed eight categories similar to those for pain and anxiety. The frequency of the categories is contained in Figure 4.3. The reliability of the codes was more than substantial with agreement ranging from  $\kappa=.81$  ( $p=.001$ ) to  $\kappa=.85$  ( $p=.001$ ).

**Figure 4.3** Category frequencies for the depression intensity descriptors.



Again, as expected *feelings/cognitions* were a category most often mentioned by participants. Almost half the references to severe depression in this category related to suicidal ideation. For example, “*life isn't worth living*” (ID 5) and “*wanting to die*” (ID 8). Others included feelings of worthlessness, isolation and extreme distress. Significant depression was

viewed less hopeless, although thoughts and feelings were still described as “*a severe mood negativity*” (ID 4) and “*you would be really sad and down*” (ID 1). “*Feeling a bit down*” (ID 12) and “*feeling frustrated*” (ID 18) were evident at both moderate and mild levels of depression.

**Intensity level**, markers or references used by participants to describe the level of the depression were often used at significant and moderate levels. Participants used the extremes of the scale (severe, mild) to anchor their judgments. For example, “*not quite as bad be as severe*” (ID 21) and “*It's not significant but it's more than mild*” (ID 5). There was a clear definition between types. In some instances severe and significant depression was described as a clinical or medical depression ( $N=4$ ). Whereas moderate depression was viewed as non-clinical or a mild clinical depression. Mild depression was not considered substantial enough to be depression ( $N=8$ ). It was described as “*feeling blue*” (ID 4) or “*feeling down not depression*” (ID 25).

An important indicator of severe depression was **behaviour** and **functioning** as indicated in Figure 4.3. Several signs of depression were mentioned, such as withdrawal, decreased appetite, fatigue, lack of care over appearance or as one participant summarised it, “*non-functioning*” (ID 4). Even at significant levels of depression there were overt indicators like tearfulness and a lack of interest in activities. Irritability and tearfulness were evident at moderate and mild levels.

As moderate depression was perceived to be controllable, **control** was a feature in participants descriptions, “*essentially manageable*” (ID 1) and “*able to cope and lift themselves out of it*” (ID 10). In contrast, inability to cope and the need for help to manage the symptom at severe and significant levels was an important feature. Most interventions were pharmacological but others were mentioned, such as “*talking to someone*” (ID 21) or hospital treatment. Essentially, significant depression was seen as more treatable than severe depression. None of the participants mentioned help with mild depression. This like anxiety was because many viewed the symptom as not depression ( $N=8$ ) or not a problem ( $N=6$ ) at that level.

All participants who made reference to **timeline** stated that this was a transient feeling of short duration at mild levels of depression. The frequency of moderate depression was more mixed with it characterised as “*just be now and again*” (ID 14) and “*a lot of the time*” (ID 26). Depression at severe and significant levels was viewed more as a bout of depression that was there all the time with no abatement ( $N=11$ ).

#### 4.6.4 Findings from the frequency descriptors

The VOICES frequency scale ranges from *most of the time*, *often*, *sometimes*, *rarely*, to *never*. Participants were asked what the midpoints in the scale meant (*often*, *sometimes*, *rarely*). These descriptions were coded in to one of seven codes:

1. All the time or almost all the time (chronic)
2. Everyday episodes (suggests frequent episodes)
3. Less than everyday but several episodes in a week
4. Once a week or less (occasional episodes)
5. Absence or near absence
6. Don't know
7. Not enough information to classify (too vague, does not answer the question).

Two researchers independently coded the transcripts. The reliability of the codes for the pain, anxiety and depression frequency descriptors was high with the strength of agreement ranging from substantial ( $\kappa = .79$ ,  $p = .001$ ) to almost perfect agreement ( $\kappa = .95$ ,  $p = .001$ ).

Code 1 (all the time or almost all the time) was the most frequent code for often: pain (13/27), anxiety (17/24) and depression (19/23). Although almost as many rated pain as code 2 (everyday episodes or frequent episodes). This suggests little differentiation between the most frequent category “*most of time*” and “*often*”. More participants found difficulty describing “*sometimes*” as 19 out of 75 replies were “*don't know*” or were too vague. Vague answers were usually a repeat of the question or did not answer the question. For example, “*sometimes is difficult judge. I expect it means different things to people*” (ID 17) or “*sometimes is I would say is just feeling sad*” (ID 25). Again, this quote illustrates an overlap with severity. Ambiguity with the descriptor “*sometimes*” was also demonstrated by the large variation in codes used. This was particularly evident for anxiety where codes ranged from code 2 (5/24), code 3 (7/24) to code 4 (4/24). For pain (13/28) and depression sometimes (17/23) most thought this was less than every day but several episodes in a week. Although a fair number (6/28) of participants thought “*sometimes*” was to have everyday episodes of pain.

“*Rarely*” was described differently depending on the symptom. For example, the majority (16/28) thought pain was once a week or less (occasional episodes), followed by absence or near absence for pain (6/28). Whereas the reverse was found for anxiety and depression, with more participants believing “*rarely*” was absence or near absence of anxiety (11/24) or depression (11/22). This was closely followed by depression (9/22) or anxiety (9/24) once a week or less (occasional episodes). Significantly, a number of participants thought that “*rarely*” was so infrequent as to not constitute anxiety (5/24) or depression (5/22). This is an important observation if the scale is a measure of these symptoms.

## 4.7 Discussion

This study revealed the complex multidimensional nature of pain, anxiety and depression. Each of the symptoms were described in terms of their psychological, social and physical impact. Participants had multiple understandings of the different terms based on their knowledge and experience. Since the interviews were conducted within the context of bereavement, these experiences were sometimes drawn upon to illustrate their descriptions. This was exemplified in one participant's reply "*after he'd gone I had these panic attacks, my legs used to go from under me... it felt like I was going to die. I would say that was severe anxiety*" (ID 7). Others used what they had observed while caring for their dying relative to explain their thoughts, "*Like my father screaming with agony I would say*" (ID 10). Despite the diversity of responses, there were clearly shared interpretations of terms. For example, in the main, severe and significant depression was considered to be a clinical level. Moderate depression was viewed as subclinical and mild, not depression at all. Similarly, mild anxiety was not considered anxiety. For pain assessment there appeared to be no problems with the descriptor "*rarely*" most proxies applying the term to pain once a week or less. A notable observation was the overlap of symptom frequency and intensity for all symptoms.

Another problematic descriptor was "*significant*"; this was for all the symptoms. Participants appeared to have difficulty expressing and/or conceptualising the term. This was evident in the number that admitted not knowing and those stating the difficulty they had describing it. The diversity of interpretations also indicated ambiguity with the descriptor. In line with several research studies, proxies were found to differ in their interpretations of vague quantifiers used to assess frequency, like "*sometimes*" and "*fairly often*" (Hammerton, 1976; Schaeffer, 1991). "*Sometimes*" had varying magnitudes; it was described as everyday episodes (suggests frequent episodes), less than everyday, but several episodes in a week or once a week or less (occasional episodes). This finding is important because terms need to be understood in a similar way if the measures are going to be used to audit care. Furthermore, it can make comparisons between ratings for the same question difficult.

An additional problem for proxies was differentiating between "*most of the time*" and "*often*". Many participants actually described "*often*" as being most of the time. This indicates that the descriptors need to be modified in order to ensure that they each have distinctive interpretations. There is evidence to suggest that individuals do not have a problem ranking the magnitude of vague quantifiers such as "*fairly often*", "*sometimes*" and "*hardly ever*" (Hammerton, 1976). This was not examined here, because proxies were already given the scale. This effectively mimics the context within which the research would be conducted. Instead, proxies were asked to quantify the pain, anxiety and depression intensity descriptors on a 0 (none) to 10 (severe) scale. Although there was some consistency in the quantities



assigned to the descriptors there were unequal intervals between the descriptors. This is consistent with the findings from Collins *et al.* (1997) who compared VAS scales with VRS in pain. They concluded that numerical indicators might be interpreted differently from VRS.

#### **4.7.1 Descriptions of the pain descriptors**

The descriptions of pain encompassed dimensions well recognised in pain research. These included sensory, affective and sensory components identified by Melzack and Torgerson (1971). The type category included sensory words such as “*stabbing*” and “*aching*”. It was found that participants had a tendency to describe lower levels of pain as an ache rather than a pain. This is consistent with cross-cultural research where the word “*ache*” has been found to denote lower intensity of pain (Gaston-Johansson *et al.*, 1990). There were also commonalities between the evaluative component and the categories intensity level and tolerance. This incorporated words such as “*excruciating*”, “*intense*” and “*unbearable*”. Tolerance, which refers to control or lack of control over the pain, was a significant category in these investigations. It was informative in distinguishing uncontrollable pain at high levels to more manageable levels of moderate and mild pain. However, perception of controllability of pain unless expressed may be difficult to assess in someone else. This may also be true for the category feelings/cognitions. Chapter 5 will examine the cues proxies use when actually assessing patients’ symptoms.

There was a definite affective component in proxies’ descriptions of pain at all levels in the category feelings/cognitions. Awareness of the pain was an important aspect of this. Affective distress from pain has been found to be an important variable in patients’ descriptions of their pain more than the sensory phenomenon (Leavitt *et al.*, 1978). Here, references to the sensory component were more frequent, probably because the participants were describing rather than actually experiencing pain. Behaviourally, this was expressed verbally or non-verbally through moaning or crying. Functioning was seen as greatly affected with a restriction in mobility and activities at significant levels. Only one participant mentioned any functioning at severe levels, which suggested that it might have been implicit that functioning was affected.

#### **4.7.2 Descriptions of the anxiety and depression descriptors**

Like pain, emotions such as anxiety and depression are multidimensional. The number of these dimensions varies across different models (Scherer, 1984; Lang, 1985). However, the most widely accepted is Lang’s (1985) three system model, which incorporates behavioural, physiological and cognitive dimensions. These dimensions were evident in proxies’ descriptions of the intensity of anxiety and depression. For anxiety, there were the physical manifes-

tations such as “*panicky*” and feelings of weakness associated with the physiological dimension of intense anxiety. There was some evidence of behavioural manifestations like irritability and outward expressions of emotions such as anger at high anxiety levels. However, this was not frequently referred to. Neither was functioning, even though anxiety can seriously disrupt activities. By contrast, these dimensions were significant at intense levels of depression, particularly for severe depression, where overt expressions of distress, in the categories functioning and behaviour, were mentioned.

As expected, the category feelings/cognitions, which in essence fits with Lang’s (1985) cognitive dimension, predominated in proxies’ interpretations at all levels of anxiety and depression. In the present study, a sense of control and help were separated out into the category tolerance. However, this is an arbitrary separation as tolerance is fundamentally cognitive. This was illustrated in one proxy’s perception of severe depression, “*things really get you down and you can’t see a way out*” (ID 21). One reason for separating out the categories feelings/cognitions and tolerance was that perception of control is only one aspect of the thoughts associated with anxiety and depression. Thus, it was possible to identify different aspects, such as attention. At intense levels of anxiety, there was an inability to focus attention on anything but fear. Whereas, at moderate and mild levels proxies believed that there was a capacity to put these thoughts out of mind. A major feature of depression was mood this was characterised by feelings of extreme sadness and negativity at intense levels. Despite these differences, there was definite overlap in proxies’ descriptions of anxiety and depression, these included irritability, crying, loss of appetite, impaired functioning and loss of concentration. This might make identifying anxiety from depression problematic especially for proxies who might not have access to the thoughts and feelings of the patient. Furthermore, overlap between anxiety and depression and the physical symptoms of the disease compound this problem (Chochinov, 2000).

Interpretation of the terms anxiety and depression can lead to variations in the reported frequency of these symptoms. Ingham and Portney (1999) point to a study where the prevalence of anxiety and nervousness varied depending on the instrument (Chang & Hwang, unpublished). In this study, many of the descriptions of intense levels of anxiety and depression were identifiable symptoms in the DSM-IV (1994) criteria for anxiety and mood disorders. Whether these would constitute clinical levels or subclinical levels is not known. Of importance when auditing the care of terminally ill patients retrospectively, is the level of symptoms associated with anxiety and depression even if they are subclinical levels. These symptoms or borderline cases of anxiety and depression are common and can have a negative impact on the patient’s quality of life (Barraclough, 1997). So, although moderate and mild

levels were not considered to be clinical anxiety or depression, they are nonetheless important in monitoring symptomology.

### 4.7.3 Limitations

The generalisability of the study must be considered. Firstly, the participants consisted of a small group of recently bereaved relatives known to the hospice services. The majority of participants had cared for their dying relative in the last week of life, although almost half of the patients had died in hospital. Their experiences may have shaped their understanding of the symptoms investigated. These experiences may not be shared by all bereaved relatives and may therefore not be representative. Secondly, culture and language are inextricably linked; the descriptors were developed by, and in this study interpreted predominantly by, individuals of caucasian western European ethnicity. The meanings attributed to the terms may be understood differently in different cultures. There is evidence to suggest that there may be differences in the types of language used to express pain in different cultural groups (Lipton & Marbach, 1984). Clearly this is an area that requires further examination.

Data from this study was coded inductively into categories so that the words, segments or paragraphs could be assigned only one code. However, in a small number of instances some words could be understood in more than one way. A good example of this was the word “*niggling*” when referred to pain, this implies a temporal element (timeline) but it may also be interpreted as sensory element (type). Similar findings have been identified with other verbal pain measures. In the MPQ several words were found to have multiple meanings. Reading *et al.* (1982) showed that “*heavy*” was in the category pressure rather than the MPQ category of “*dullness*”. Ideally, one method to overcome the ambiguity in the present study would be to ask the participant what they meant by the word. Given the context of the study and the practicality, this was rejected. Instead, where these instances arose the word was interpreted within the context of the sentence by two independent coders. Agreement was reached on whether the word was coded as timeline or type.

## 4.8 Conclusion

To date, there is no information on the reliability and validity of VOICES. This is currently under investigation by the Department of Palliative Care and Policy at King’s College, London. The main objective here was to elicit proxies’ interpretations of descriptors used in VOICES as a means to improving the instrument. From the findings, it is clear that some terms need to be modified to ease clarity for proxies. Further testing would be necessary to examine whether these modifications reduced ambiguity. These types of investigations have an important role in survey design, as the designer’s meaning may not be consistent with the



participants. For example, “*significant*” is a clinical term, it is possible that participants had no understanding of the term in this context to explain symptoms. Despite the fact that VOICES is a unidimensional scale participants’ perceptions revealed that the symptoms of pain, anxiety and depression could be understood within this in multidimensional ways. A problem with this is that there is no way of disentangling particular aspects of the symptom, for example, affective components from sensory components of pain. However, these types of unidimensional scales may be more appropriate retrospectively as they are less burdening for proxies than multidimensional scales. Moreover, unidimensional scales might be all that is required to identify where patients’ needs are not met and for informing health care providers on the quality of service. In Chapter 5, proxies will use VOICES to assess their relative’s care.

# CHAPTER 5: Proxies' symptom judgments

## 5.1 Introduction

To date, research on the validity of proxies for patients at the end of life has focused on comparing patients' accounts with those of their proxy. The review (Chapter 2) showed there are definitely inconsistencies in proxies' perception of particular aspects of the patient's experience. Further still it highlighted the need to go beyond comparisons between the patient and proxy. An understanding of how proxies answer questions regarding the patient's care in a retrospective context might expose difficulties and possible limitations of their use in palliative care research. This includes an investigation of proxies comprehension of the questions being asked, how they retrieve information and make judgments. It is also useful to be aware of the cognitive processes used by proxies as it may help reduce errors when designing questionnaires. This chapter investigates what proxies are recalling when they report on patients' levels of pain, anxiety and depression. This is a particularly important issue because awareness of the possible biases and errors will allow investigators to identify cues used by proxies and to interpret the basis of their judgments. Although the research presented here is based on the VOICES survey, the results are applicable to the retrospective approach in general and may be useful in identifying ways to improve existing survey designs.

## 5.2 The cognitive model of question answering and proxies

Survey questions rely heavily on cognitive processes and memory. In order to understand the position of the proxy in the context of retrospective palliative care studies, the emphasis of the present investigation is on the cognitive model of question answering (detailed in Section 3.2). Errors can occur at any stage in this process. As discussed in Chapter 3, using the cognitive interview as a technique to assess proxies' comprehension, memory and decision-making, it is possible to analyse the quality of the data provided (discussed in Chapter 3). Areas of interest are response difficulties, cues to the symptom, frequency judgments, and the reference periods assessed. Before describing the study, it is necessary to outline research from survey methodology and cognitive psychology as it applies to these investigations.

### 5.2.1 Response difficulties

Coding response difficulties are an important process in survey design, as a means of pre-testing questions and evaluating procedures. In Chapter 4, an investigation of the symptom descriptors from VOICES helped reveal proxies' interpretations of terms and symptoms. This was an important step because it exposed ambiguities and divergent interpretations of terms, both of which can impede attempts to obtain valid responses. Proxies may also encounter dif-

difficulties answering questions during bereavement, another critical issue requiring some prior exploration. Within survey methodology, this process usually involves monitoring the participant and the interviewer to identify problems with questions, a technique known as *behavior coding* (Fowler & Cannell, 1996). Examples include inadequate answers, “*don’t know*” responses, requests for clarification and uncertainty about the accuracy of a response. In previous research, behaviour codes such as these have been useful in diagnosing problems with questions (Oksenberg *et al.*, 1991; Cannell *et al.*, 1992). Cannell *et al.* (1992), in an analysis of response difficulties across seven surveys, identified common problems associated with specific codes. For example, questions that needed clarification are those that often require means or information that are outside the participants’ experiences or frame of reference. Inadequate responses, on the other hand, are indicative of a level of detail that the participant is not able to provide.

### **5.2.2 Memory for events**

Following comprehension of a question, the participant has to retrieve the relevant information from memory. As discussed in Chapter 2, errors can occur for a number of reasons. The time interval between encoding and retrieval is one of the main factors affecting the recollection of events accurately. The greater the interval the more likely it is that an error will occur (Baddley, 1990). The exact relationship itself is not straightforward and varies according to the frequency and saliency of the events in question and what occurs in the intervening period. Salient events are those that are unique, emotional and mark a transition in the participant’s life (Linton, 1982). Such events can be more easily and precisely recalled (Bower *et al.*, 1979; Bower, 1981). However, the salience of an event can lead to over-reporting (Huber & Power, 1985) because the ease of recall may lead the participant to believe that event has occurred more often or more recently in time (Tversky & Kahneman, 1973; Brown *et al.*, 1986). This may cause more frequent, yet less salient events to diminish in importance. They may become less distinguishable and forgotten or merged into a generic memory that represents a summary of events (Linton, 1982). There exist a substantial body of research on the effects of the individual’s emotions on memory at the time of the event and at recall. This is an area that will be discussed in greater detail in a Chapter 6.

### **5.2.3 Cues and judgments**

In these investigations, knowledge of what cues proxies use to devise their responses is useful for understanding the validity of their responses and for the management of symptoms. For example, are proxy judgments based on observations, communication with the patient, or on assumptions about the patient’s experiences? In Chapter 4, proxies understanding of the

symptom descriptors was explored. This is a useful framework within which to examine how proxies' arrive at their answer. There is a paucity of literature in this area, and what has been conducted suggests that cues have different effects on symptom reporting. In some instances, reliance on verbal cues may lead to inaccuracies because of denial or a need to protect significant others (Lobchuk & Kristjanson, 1997). This may be particularly marked for symptoms that are less visible where there is a lack of overt behavioural cues. Whereas, other more overt manifestations may provide proxies with reliable cues to patients' experiences, and could explain congruence between patients and their proxies.

It is important to consider judgments within a social context. Responses are liable to be based on the proxy's general knowledge, beliefs and expectations. For this, it is necessary to turn to the literature on self-knowledge as a means of understanding the basis of responses (Klein & Loftus, 1993; Srull & Wyer, 1989). Although much of the work has centred on self-knowledge, several experiments have been conducted on knowledge of others to examine whether there is a difference (Klein & Loftus, 1993). The findings indicate that trait knowledge (knowledge based on the person's characteristics/personality) of others, founded on past behaviours, is likely to be accessed by participants when answering questions (Schwarz & Sudman, 1996). Evidence to support this comes from Blair *et al.* (1991) who found that proxies were more likely to rely on knowledge about the person they were reporting for than when self-reporting. In the present study, it might be that proxies use trait knowledge to judge the patients coping with symptoms. Thus, trait knowledge may form the basis for judgments about symptom presence and severity.

Normative behaviour and general knowledge (including trait knowledge) are two sources of knowledge used by participants identified in the literature (Schwarz, 1990). Normative behaviour within the context of palliative care might include a comparison of the patient's symptoms with other patients or expectations about what is experienced during terminal illness. General knowledge might include comparisons with self or over time, or as already mentioned, it may be based on patients' traits. When proxies do not have all the relevant knowledge to make a judgment, they may rely on heuristics (e.g. rule-of-thumb methods) to reach a decision (Kahneman & Tversky, 1971; 1973; Tversky & Kahneman, 1974). For instance, they may anchor their judgment on norms or general knowledge and adjust their response accordingly; the anchor and adjustment heuristic. Alternatively, they may choose a response that appears representative of the population being studied; the representativeness heuristic. These heuristics may bias judgments and lead to inaccuracies in reporting (Tversky & Kahneman, 1974; Kahneman *et al.*, 1982).

### 5.2.4 Frequency judgments

As Willis and colleagues' (1991) model shows (Section 3.2), information is retrieved from memory and the adequacy of the information is assessed in relation to the question. Much of the work in survey methodology has focused on frequency judgments and error as a result of the strategies individuals use. This type of analysis is pertinent to the present investigations, as proxies are asked to gauge the frequency and severity of symptoms. There are a number of well-recognised strategies in the cognitive and survey methodology literature, among the most commonly identified are episode enumeration, rate-based estimation and heuristic-based estimation (Tversky & Kahneman, 1973; Sudman & Bradburn, 1974; Kahneman *et al.*, 1982; Blair *et al.*, 1991; Bickart *et al.*, 1994). Episode enumeration is the recall and count of each episode within a specified period and is considered generally more accurate than estimation (Clark & Schober, 1994). Although recall and counts are liable to errors of omission due to forgetting, or commission where events outside the reference period are recalled (Sudman & Bradburn (1973), the accuracy and use of this method has been found to vary as a function of the length of period being assessed and the frequency of the occurrence (Burton & Blair, 1991). For example, episode enumeration is unlikely over periods of months, due to recall errors and the degree of cognitive effort required. In these instances participants may use strategies to estimate frequency instead.

Rate-based estimations are a strategy whereby the participant assesses the frequency over a shorter period and multiplies the frequency over the period being assessed (Bradburn *et al.*, 1987). Precision relies on the regularity of occurrence. For example, "*I visit the doctor every month, so that is twelve times in a year*". This may improve accuracy for regular events but is likely to be less accurate for invariable occurrences. There is less emphasis in the survey methodology literature on heuristic-based estimations. However, as mentioned earlier, there is considerable literature in cognitive psychology dedicated to the strategies individuals use in reaching decisions (Tversky & Kahneman, 1973; Kahneman *et al.*, 1982). A notable example is the *availability heuristic*, which states that the likelihood of an event having happened is judged by the ease with which events can be recalled (Tversky & Kahneman, 1973). As Bradburn *et al.* (1987) points out, this is the cause of many errors in a number of surveys. Based on the availability heuristic, Eisenhower *et al.* (1991) predict that the direction of bias should be an overestimation of salient events and an underestimation of frequent events. Although the present study will not test this hypothesis, an exploration of proxies' cognitive processing will nonetheless increase awareness of how inferences are made.

### 5.2.5 Reference periods and event cues

Surveys usually require proxies to recall events within a specified reference period. There is no doubt that the death of a significant other is a vivid and emotional event that is easily brought to mind because of its importance. However, periods leading up to the death may be more memorable than others because of their salience, and as a result of their timing in the sequence of events. For example, cues such as social occasions, events of personal significance and calendar dates can prompt recall and improve accuracy (Bradburn *et al.*, 1979). This is because memory is organised into sequences of events within a personal time frame (Brown *et al.*, 1986). Based on these findings, it might be easier for proxies to recall what happened when, for instance, the patient was admitted to the hospice, rather than what happened over the last year or month. Another common problem with arbitrary reference periods (week, month, year) is *forward telescoping*. This is when events outside the reference period are imported into it; as a consequence frequencies are overestimated (Sudman & Bradburn, 1973). One explanation for this is that the clarity or saliency of an event is thought to lead the individual to the belief that the event occurred more recently than it actually did (Brown *et al.*, 1986). The significance of the information sought in the present investigations may be particularly prone to this problem. To some extent, forward telescoping may be responsible for proxies' overestimations of patients' pain, anxiety and depression observed in several studies (O'Brien & Francis, 1988; Higginson *et al.*, 1990; Higginson & McCarthy, 1993; Spiller & Alexander, 1993; Field *et al.*, 1995; Yeager *et al.*, 1995; Elliott *et al.*, 1996; Sneeuw *et al.*, 1997b; 1998; 1999).

The length of the reference period and the content of the material being assessed are important factors in the strategies used by participants. By way of illustration, VOICES has a twelve month reference period (Addinton-Hall *et al.*, 1998). For the evaluation of some aspects of care at the end of life, such as services received, this may be appropriate. However, for symptoms with variable frequencies and intensities, this period may be too long and is likely to result in inaccuracies. This is because the recall of episodes over such an extended period is improbable and requires enormous cognitive effort (Blair & Burton, 1987). The variable nature of symptoms does not lend itself to rate-based estimations, thus proxies are more likely to rely on estimation heuristics. Estimations may be biased due to the limitations of memory. For example, events nearing the patient's death may take precedence because they are more recent and accessible in memory. A consequence of this is that proxies may be reporting on a small interval within the actual twelve month reference period. Thus, what appear to be arbitrary reference periods have an enormous impact on the accuracy of survey data.

## 5.3 Aims

The study carried out in this Ph.D. investigates what proxies are recalling when they report on patients' levels of pain, anxiety and depression. The aim of this study is to identify how proxies decide what answer to give when asked about the patient's pain, anxiety and depression. The objectives are to identify (1) response difficulties (2) strategies upon which decisions are made, (3) cues used to form the basis for symptom judgments (4) reference periods. Greater knowledge of these processes will help in understanding possible biases and errors when using proxies for patients at the end of life.

### *Research question two*

What are proxies recalling when they report on patients' levels of pain, depression and anxiety retrospectively and what cues they are using?

## 5.4 Methods

As detailed in Section 3.2, the methods and procedures included cognitive interviewing using retrospective verbal protocol. Following questions about the presence, frequency and severity of patient's pain, anxiety and depression proxies were asked open-ended questions requesting them to verbalise their thought processes. For example, after answering a question regarding patient's pain in the last week, proxies are asked, "*Can you tell me what were you thinking about when you answered that question?*". The cognitive questions directly followed the VOICES question and are a means to discover how proxies arrived at their answer. The Proxy Semi-structured Cognitive Interview (PSCI) was developed by the investigator as a framework within which to organise and obtain this information from proxies (Appendix F).

A coding scheme was used as a framework to deductively assign codes to the transcripts. This framework was based on a coding scheme developed by Sudman *et al.* (1996), with a theoretical scheme incorporating comprehension, retrieval and judgment. The codes used for identifying response difficulties were based on a protocol devised by Fowler and Cannell (1996). As the analysis also incorporated codes that emerged from the data through an inductive content analysis, the PSCI coding scheme will be discussed along with the analysis in the next section.

For the reasons identified in Section 5.2.5, and because patients are more likely to rely on proxies as their health deteriorates near death, the last week and month of life were chosen as reference periods for the present investigations.



## 5.5 Findings and discussion

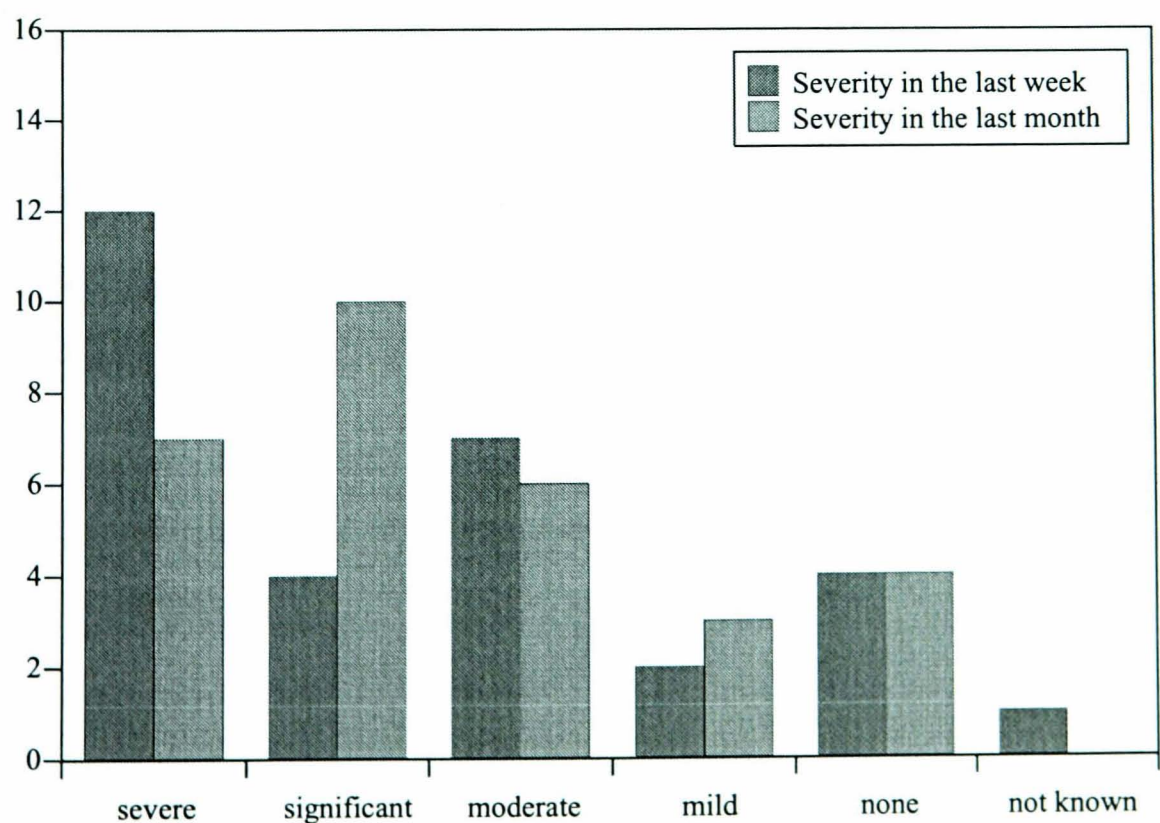
Descriptive statistics of patient and proxy characteristics are summarised in Tables 4.1 and 4.2. Proxies' responses to the VOICES questions and numerical response data were analysed using SPSS/Win 10.1. Content analysis of the PSCI was performed through winMAX 98Pro, a qualitative analysis program and Microsoft Excel spread sheets. The first part of this section of the chapter contains descriptive statistics of proxies ratings of patients symptoms using VOICES. This is followed by analysis of the PSCI.

### 5.5.1 Proxies' responses to VOICES

#### *Pain*

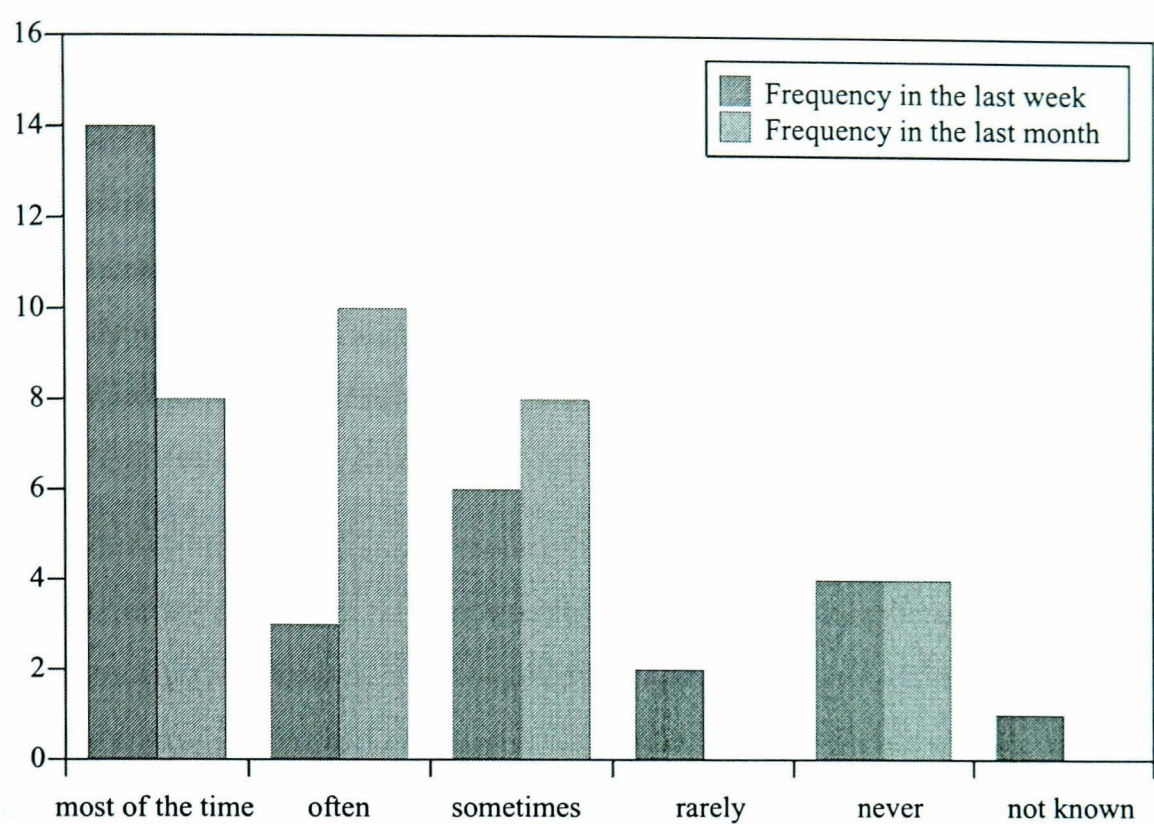
Proxies' ratings of patients' levels of pain severity and frequency (Figure 5.1 & 5.2) revealed that they perceived their relative to be in a considerable amount of pain in both the month and week before their relative's death. Three quarters of the sample stated that the patient's pain was moderate or greater during both times. The tendency was for the degree of pain to be rated more severely during the last week of life. Almost half of the sample rated the severity of pain as severe in the last week suggesting that they believed the pain was particularly acute during this period. The frequency of pain was also rated as more frequent during the last week of life with 14/30 stating that their relative was in pain most of the time compared with 8/30 in the last month.

**Figure 5.1** Proxies' ratings of patients' pain severity in the last week and month of life.





**Figure 5.2** Proxies' ratings of patients' pain frequency in the last week and month of life



Twenty-eight patients were reportedly taking analgesia for pain. Of these, nineteen patients got some or partial relief from the analgesia. A further 6/28 believed that the analgesia was totally effective in relieving their relative's pain. Two proxies were unable to determine the analgesic effectiveness, while one proxy stated that the analgesia was not effective at all. In the vast majority of cases, the hospice (13/28) provided patient's pain relief. Followed by their GP (6/28), hospital (4/28) or a combination of sources.

**Anxiety**

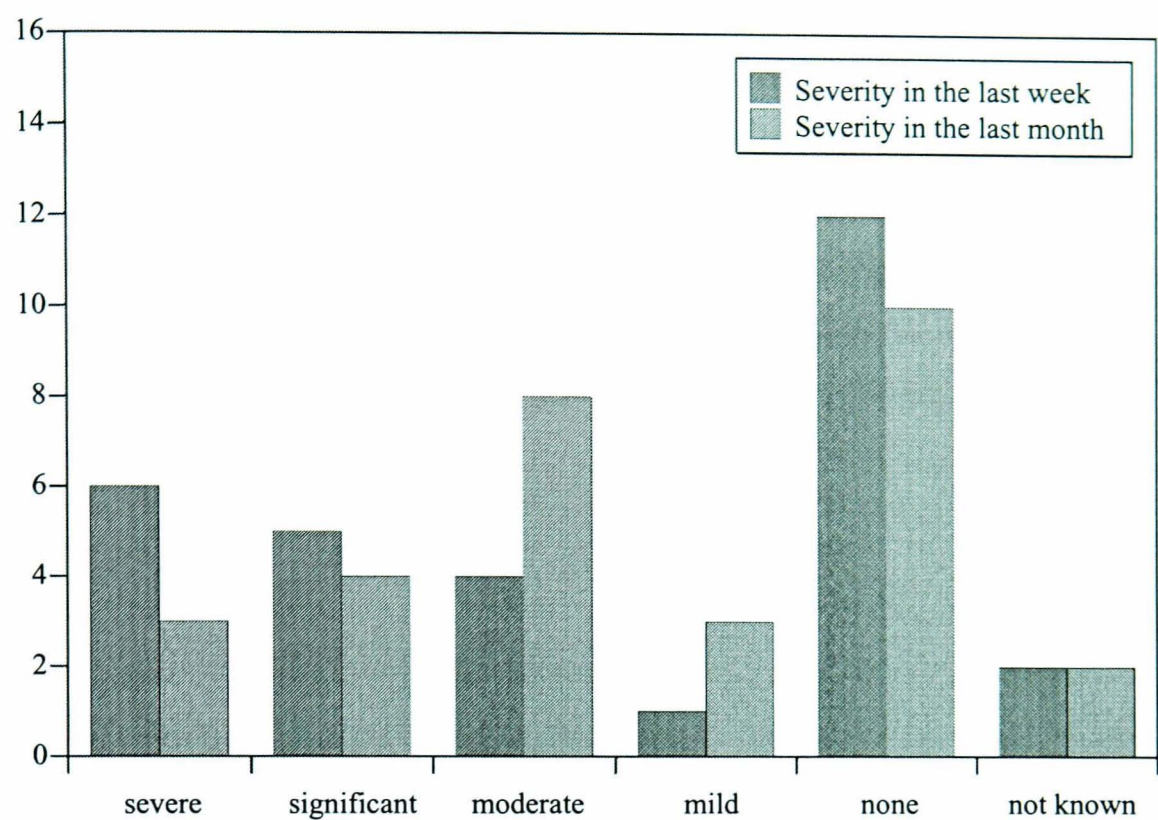
Proxies' ratings of patient's anxiety (Figures 5.3 & 5.4) suggested that many believed that anxiety was not present. Despite this finding, a substantial proportion thought that their relative's anxiety was moderate or greater in the last week (15/30) and month (15/30) before death. Few proxies reported that their relative had received any help to cope with their anxiety (3/30). Medication in the form of tranquillisers (2/30) or counselling (1/30) were the types of help given. These were thought to alleviate the anxiety temporarily in two patients.

**Depression**

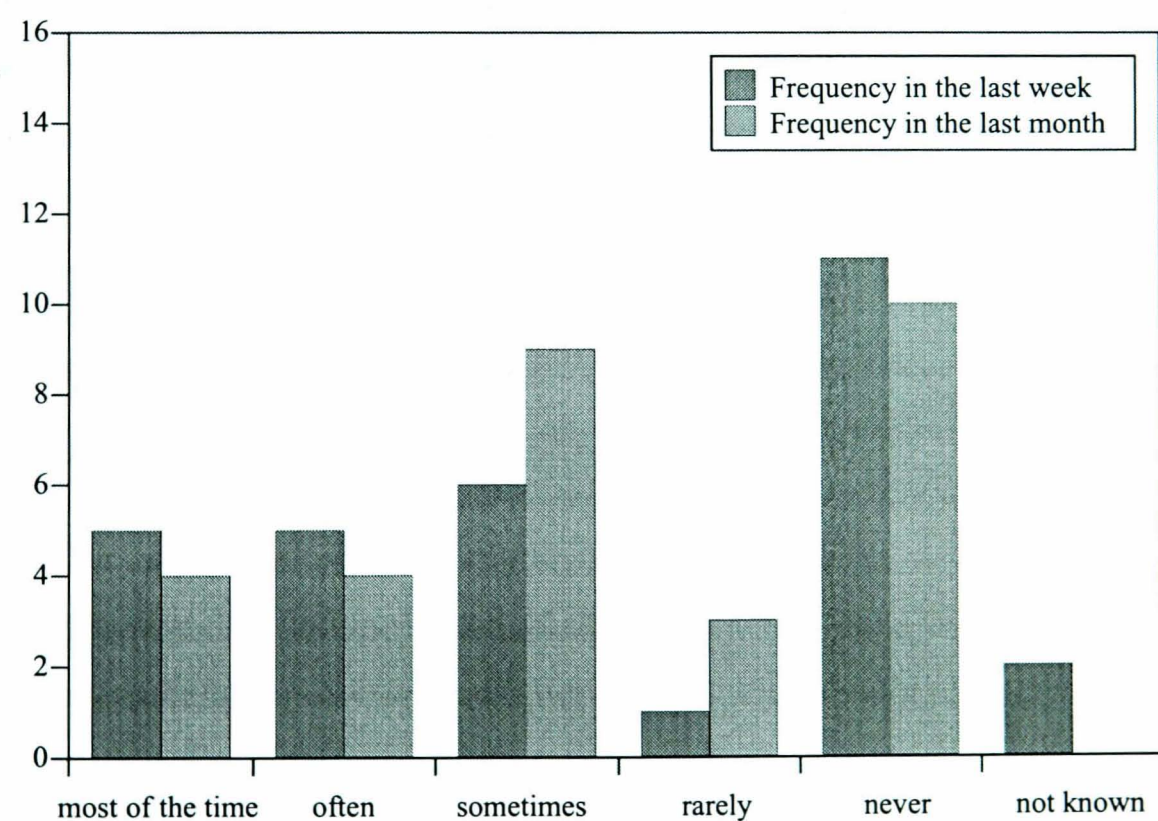
A notable observation from proxies' ratings of patients' depression (Figure 5.5 & 5.6) was that almost half of the sample reported no depression in either the last week or last month of life. Furthermore, there were more patients rated as having levels of depression that were of moderate or greater severity in the month (14/30) than in the week (12/30) before death. A possible reason for this might have been that proxies appeared to have greater difficulty



**Figure 5.3** Proxies' ratings of patients' anxiety severity in the last week and month of lif

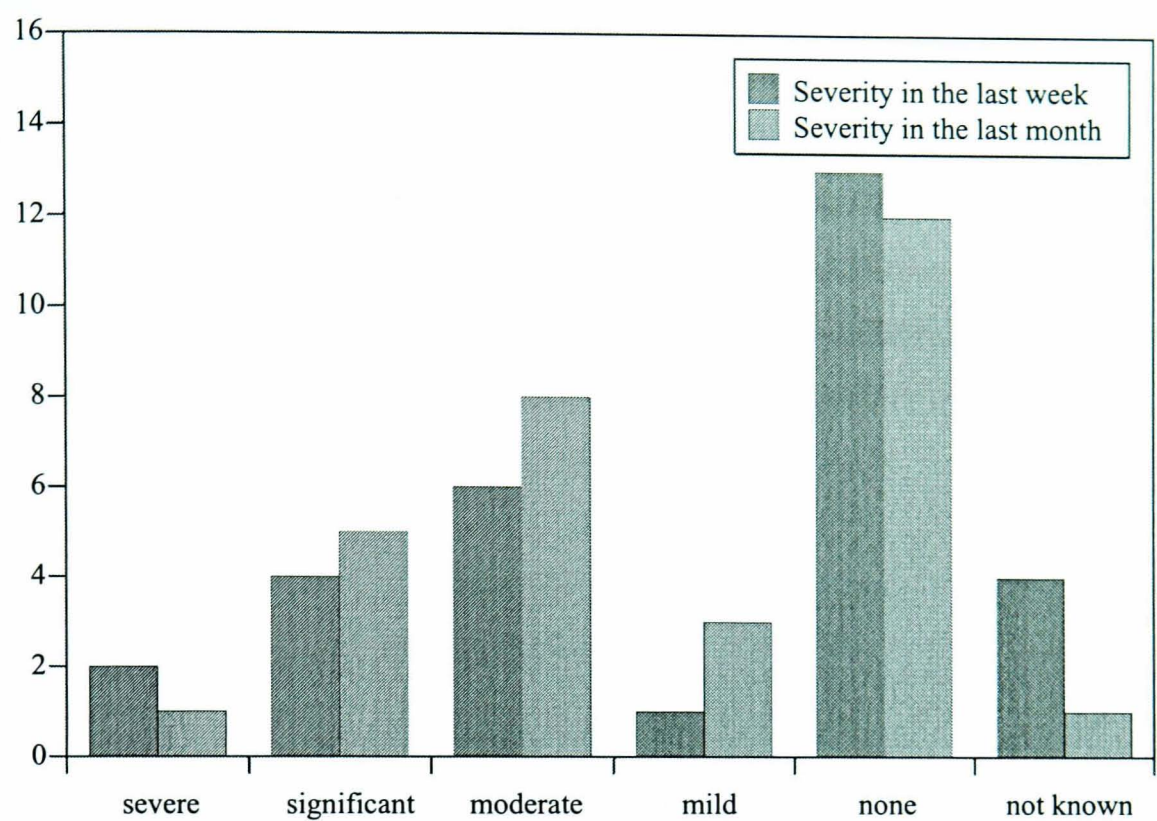


**Figure 5.4** Proxies' ratings of patients' anxiety frequency in the last week and month of l

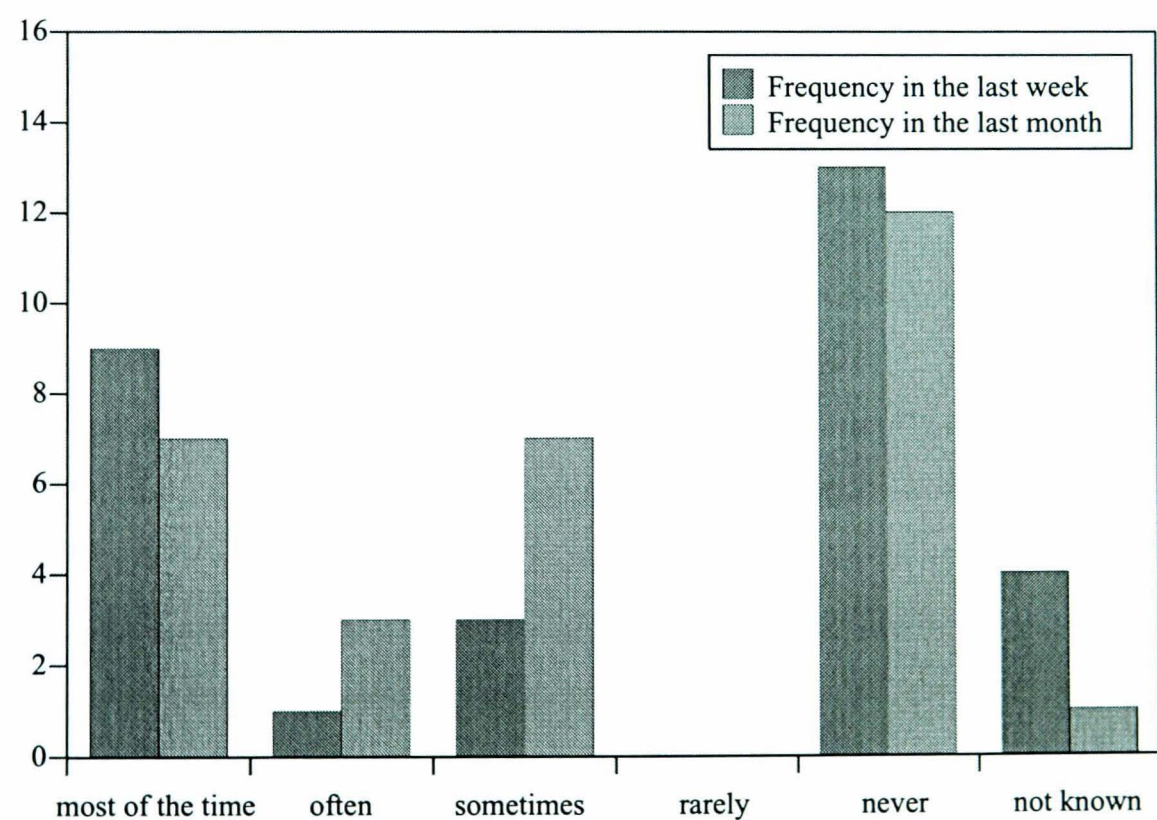


deciding whether their relative had depression in the last week (4/30) than in the last month (1/30). Six patients were reportedly receiving help to cope with their depression. Five were taking medication and one patient was receiving counselling. The counselling was thought to be effective sometimes, while proxies' viewed the medication interventions as either ineffective or providing minimal relief.

**Figure 5.5** Proxies' ratings of patients' depression severity in the last week and month of life.



**Figure 5.6** Proxies' ratings of patients' depression frequency in the last week and month of life.



**5.5.2 Analysis of the PSCI**

Here, the emphasis was on what proxies were recalling when they reported on patients' levels of pain, depression and anxiety retrospectively. An inductive and deductive content analysis of proxies' verbalisations was performed. A framework to deductively assign codes to the



transcripts was based on previous coding schemes (Fowler & Cannell, 1996; Sudman *et al.*, 1996) An inductive content analysis was also used to identify themes that emerged from the data. Table 5.1 contains the PSCI coding scheme. Codes identified by the investigator and specific to these investigations are indicated in the coding scheme with an asterix (\*).

The process began by coding the first three transcripts. The unit of analysis was the response or segments of the response. This involved identifying key words and themes within the text. For example, response difficulties, cues to symptoms, strategies for estimating symptom frequency and severity, reference periods and event cues. As new theme emerged, these were integrated into the coding system and used as a framework for further analysis of the transcripts. Codes varied depending on the type of question asked. A case in point was the symptom frequency codes. These were designed specifically to examine how proxies reached a decision on the frequency of a symptom, so were not applied to other questions.

**Table 5.1: Coding scheme for PSCI.**

|   | Code description  |
|---|---|
|   | <i><b>Response difficulties</b></i>   |
|   | No verbalisation of strategy  |
|   | No apparent strategy ( <i>e.g. guess</i> )  |
|   | Don't know  |
|   | Incoherent response   |
|   | Explicitly states that the question is difficult to answer no reason given why  |
| * | States that the question is difficult to answer because of patient's level of consciousness ( <i>e.g. due to medication or confusion</i> )  |
| * | States that the question is difficult to answer because of lack of communication between proxy and patient ( <i>not due to patient's condition</i> )  |
| * | States that the question is difficult to answer because of overlap with pain, anxiety and depression  |
| * | Unable to give a response ( <i>e.g. too distressed or makes reference to the fact that they do not want to think about it</i> )   |
|   | Inadequate response ( <i>e.g. response does not answer the question, talks about another symptom</i> )  |
|   | Expresses uncertainty ( <i>e.g. words such as "probably" or "doubt", or asks for clarification</i> )  |
| * | Response changed during reply ( <i>e.g. says "severe" then changes to "significant", or says "no" to symptom then changes to "yes" during response</i> )  |
|   | <i><b>Symptom frequency judgments only</b></i>  |
| * | Symptom constant/almost constant (no frequency estimation)  |
|   | Episode enumeration<br><i>Here the proxy will recall particular episodes within the period of assessment. Their judgment will be based on a count each of the episodes (e.g. "He had pain on the Wednesday and on the Friday in the week before he died")</i> |

Table 5.1: Coding scheme for PSCI.

| Code description                 |   |
|----------------------------------|---|
|                                  | Rate-based estimation<br><i>Here the proxy will base their judgment on the regularity of the occurrence of the symptom (e.g. "She normally got the pain every night")</i>   |
|                                  | Enumeration estimation<br><i>Here the proxy will base their judgment on the variable number of occurrences of the symptom. This would include strategies such as averaging over the assessment period or other strategies excluding rated-based estimation and counting/recall episodes (e.g. "It wasn't as often at the beginning but it gradually became worse so over that month I would say often")</i> |
| <b>Comparative judgments</b>     |   |
|                                  | <i>Anchoring</i>  |
|                                  | Anchor on norm what would be expected (e.g. "I would think most people would be depressed if they knew they were going to die")   |
|                                  | Anchor on self or another (e.g. "he was better at coping with it than me")  |
|                                  | General characteristic/trait of the patient (e.g. "He had always been a positive person")   |
| *                                | Anchor on what they were like previously during the illness (e.g. "he was worse in the last week than in the month")  |
|                                  | Anchor on another response option and adjust (e.g. "It wasn't severe it wasn't as bad as that so I would say significant" or "It was more often than sometimes so I would say often")   |
|                                  | <i>Intensity rating questions only</i>  |
| *                                | Variable symptom severity with rating based on the most frequent rating   |
| *                                | Variable symptom severity with rating based on the most severe rating   |
| *                                | Variable symptom severity with no explanation for choice of rating  |
| *                                | Intensity rating based on average of variable pain intensity (e.g. "it was severe in the last week, but the week before it was mild, so I would say overall it was moderate")   |
| *                                | Frequency used as an indicator of symptom severity  |
| <b>Basis of symptom judgment</b> |   |
| *                                | Reference to verbal communication between patient and proxy   |
| *                                | Reference to non-verbal communication, behaviours or cues to symptom such as the patient's appearance   |
| *                                | Reference to verbal communication with a third party (e.g. nurses, family members)  |
| *                                | Reference to patient's thoughts and feelings (e.g. frustration, concentration, anger)   |
| *                                | Reference to patient's functioning (e.g. activities of daily living, social functioning)  |
| *                                | Reference to patient's coping/control (e.g. references to control or lack of control, coping)   |
| *                                | Reference to medication/help (e.g. explicit references indicating help with the symptom)  |
| <b>Event cues</b>                |   |
|                                  | Person mentioned (e.g. nurses, family)  |
|                                  | Place mentioned (e.g. hospice, home)  |
|                                  | Characteristic of event mentioned (e.g. last rites, dressing changes, etc.)   |

Table 5.1: Coding scheme for PSCI.

| Code description                           |  |
|--|--|
| <i>Reference period</i>                    |  |
| *  | Anchor on a significant event ( <i>e.g. illness, hospital admission</i> )                                    |
|  | Anchor on a date ( <i>e.g. wedding, holiday</i> )  |
|  | Anchor on a season of the year ( <i>e.g. summer, spring</i> )  |
| *  | Mentions reference period being assessed ( <i>e.g. week or month</i> )                                       |
| *  | Reference to period outside the specified period ( <i>last month or week</i> )                               |
| <i>Proxies' emotions</i> <sup>†</sup>      |  |
| *  | Proxy's emotional reaction to event mentioned  |
| <i>Affect-based judgments</i> <sup>†</sup> |  |
| *  | Explicit inferences about how the patient was feeling or thinking based on proxies own feelings about events |

† Proxies' emotions and affect-based judgement are analysed and discussed in Chapter 6.

5.5.3 Reliability/validity

Since the overall validity of the methodology was discussed in Section 3.3.2, the focus now moves to the reliability and validity of the interpretations of the individual transcripts. To ensure that the analysis provided a valid interpretation of what proxies were saying, the comprehensiveness and appropriateness of the analysis was discussed with another researcher. The codes were then applied and compared across transcripts during the analysis in order to assess the validity and consistency of their interpretation. To fully capture the meanings in what proxies were saying, additional codes were added to the framework as they emerged from the transcripts.

Reliability was evaluated through consistency across the transcripts and between raters. Two coders independently coded the transcripts in order to examine the reliability of the analysis. Cohen's kappa was used to evaluate the level of agreement. This is a more comprehensive measure of inter-rater agreement because, unlike percentage agreement, it takes into account chance agreement. Kappa ranged from .80 ( $p < .01$ ) to .93 ( $p < .01$ ). Mean kappa values for each of the twenty-one questions was .88 ( $p < .01$ ) (s.d. .03). According to Landis & Koch (1977) these estimates suggest substantial to almost perfect agreement between coders. Where differences between the two raters emerged, the divergent interpretations of the code or response in question were discussed and agreement was reached. Although it might appear attractive to compare kappa values across studies, the measure itself is only meaningful if the number of categories used to estimate agreement is constant (McClure & Willett, 1987). However, Sudman *et al.* (1996) using a similar framework report inter-rater reliability levels

of more than .80 for broad categories and lower levels for finer categories. Code frequencies and kappa values for all the questions are contained in Appendix N.

### 5.5.4 Response difficulties

Since few of these codes were applied to the transcripts (see Appendix N), it would appear that proxies encountered few of these types of response difficulties. Amongst these were inadequate responses, no apparent strategy and incoherent responses. One reason for this might have been that response difficulties were specific to the area of investigation. As Table 5.1 shows, proxies had trouble answering questions relating to patients' symptoms for a number of reasons. These included a lack of communication between the patient and proxy, an overlap in symptoms, or because of the patient's decreased level of consciousness. This was particularly true when the proxy was asked to identify the presence of anxiety (6/30), depression (5/30) and pain presence (3/30) and severity (4/25) in the last week of life. As one participant remarked,

*"She was unconscious so it was difficult to tell at that stage. She might have been in pain but not able to communicate."*

(ID 1; 3b)

Estimates suggest that sedation of patients nearing death, through narcotics or psychotropic drugs is not uncommon in controlling some symptoms, such as pain. Levels ranging from 16% to 52.5% have been reported (Ventafridda *et al.*, 1990; Fainsinger *et al.*, 1991). This poses considerable difficulties for proxies when assessing patients and present a challenge for researchers using the retrospective approach.

A further difficulty in assessing patients' symptoms was a lack of communication between the patient and proxy. This was a notable problem when detecting the presence of depression (4/30) and anxiety (3/30) in the month reference period, and anxiety in the week reference period (3/30). This was evidenced in a response about the presence of anxiety,

*"I can't be completely certain .. there were no verbal signs nothing to indicate that."*

(ID 18; 11b)

These problems were often associated with uncertainty regarding the severity of these symptoms; particularly for depression. Disruptions in communication between patients and their partners are not unusual as they try to cope with the patient's illness and impending death (Baider & Sarell, 1984). Furthermore, coping strategies such as denial may inhibit the open expression of feelings between the patient and proxy. Unless these symptoms are communicated to the patient via verbal or non-verbal means, the proxy will have no way of knowing how the patient is feeling. References were also made to the effect that patients might have been concealing their feelings so as not to upset or burden significant others. There is some

evidence to support this finding. Dar *et al.* (1992) found that 60% of patients admitted hiding their pain in order to avoid upsetting their spouses. Psychological symptoms, like anxiety and depression, may be easier to conceal than more overt symptoms such as breathlessness. This could lead to an underestimation of patients' symptoms. A finding observed in Hinton (1996) and Higginson *et al.*, (1994) retrospective observations, where anxiety and depression were rated less severely. However, it was noted that some proxies bear this in mind and infer from the patient's character and the situation when estimating anxiety and depression, as one participant stated of his wife's depression,

*"Although she was good at hiding the anxiety and that, she did want to talk to us to say her goodbyes and that in the last week. And so, you know, that obviously must have been sad for her."*

(ID 6; 20b)

Consequently, some proxies may actually be more in line with the patient's rating, as Hinton (1996) found for anxiety rated retrospectively. On the other hand they may overestimate anxiety and depression, a bias consistently observed in prospective studies comparing patients and proxies estimations of patients' psychological status (Higginson *et al.*, 1990; Higginson & McCarthy, 1993; Spiller & Alexander, 1993; Field *et al.*, 1995; Sneeuw *et al.*, 1997b; 1999).

A small number of proxies (3/30) were not able to decide whether the patient had depression or anxiety in the week that they died because of the overlap between the symptoms of depression, the patient's physical condition, and anxiety. The problems were summed up in a response about the presence of depression,

*"I couldn't say he had depression. It was difficult to tell because it was. Well it seemed to be mixed in with the anxiety to some extent."*

(ID 4; 24b)

As mentioned in Section 4.1.2, symptoms such as fatigue, weight loss and poor concentration are common in patients with cancer (Chochinov, 2000). In addition, performance status is known to decline near death (Vigano *et al.*, 2000). Altogether, these features often overlap with the symptoms of depression and anxiety. These include sleep disturbance, fatigue, difficulty concentrating and significant distress or impairment of social, occupational or other areas of functioning (American Psychiatric Association, DSM-IV, 1994). Since proxies rely on these types of overt expressions when deciding whether the patient is anxious or depressed, any overlap with their physical condition can lead to significant problems identifying depression and anxiety.

Of those proxies who reported pain in the last week (25/30) and month (26/30), more difficulties were experienced when deciding on the frequency and severity of pain in the week rather than in the month. By contrast, a similar number of problems were encountered with



anxiety and depression severity and frequency in the last week and month of life. For all symptoms though, more response difficulties were evident for symptom severity than for frequency. There were almost double the numbers of problems for depression severity compared with depression frequency. It is not known why this occurs, although it may be possible to conjecture that the frequency of a symptom can be to some extent observed, whereas the intensity of the symptom is largely a perceptual experience, thus less discernible to the onlooker.

### 5.5.5 Symptom frequency judgments

Symptom frequency judgments require the proxy to retrieve information from memory, evaluate this material and formulate a response that meets the requirements of the question. As stated, errors can occur at any time in this process. In these investigations few stated that they could not recall the information requested. This was probably due to the length of the reference periods and the saliency of the memories surrounding the death of their significant other.

#### *Pain*

In this study, proxies used a range of strategies to judge the frequency of patients' symptoms. This varied according to the symptom and the period being assessed. Of those who reported pain a number stated that the patient had pain all the time during the last month (8/24) and week of life (8/25), implying that there was no estimation of frequency in these cases. Few used episode enumeration in either the week (4/25) or the month (4/24) assessment periods. In the last week, episode enumeration was used for episodes of pain that were less frequent; "rarely" or "sometimes". This is consistent with the literature where the strategy is more likely to be used when there is a short assessment period and infrequent episodes (Blair & Burton, 1987). However, during the month assessment period, recall and count of episodes was used for pain reported as "often." In both of these cases, the participant recalled each episode when they had visited the patient and remembered the patient being in pain on each of these occasions. In general, for frequent episodes of pain, almost equal numbers used either rate-based estimation or heuristic based estimations. Rate-based estimations tended to be used when there was some regularity in the analgesic regime or the pain was more salient at particular times (such as during the night). Heuristic based estimations characterised the variability of the pain experienced. This was exemplified in one participant's reply,

*"It got worse you know progressively worse over that month .. He was getting about then and even getting in and out of bed at one point, but then it got worse towards the last week or two...within two weeks he wasn't getting up."*

(ID 28; 9b)

## **Anxiety**

Of the proxies who reported anxiety, almost half (8/16) recalled and counted actual episodes during the last week of the patient's life and six out of twenty proxies in the month. This strategy was evident for frequent and infrequent episodes of anxiety. One reason for this might have been where anxiety was reported as "*often*" or "*most of the time*", the episodes were not discrete short periods of anxiety but over prolonged periods. This contrasts with other survey appraisals of frequency where there are discrete episodes such as doctor's visits, or alcohol and food consumption (Blair *et al.*, 1991; Armstrong *et al.*, 2000). In the present investigations an event where the patient became anxious such as during the last rites or a dressing change, could lead to a prolonged episode of anxiety that could last days. Rate-based estimations were seldom used in either the month or week. This might have been due to the variable nature of anxiety. In the last month, over half (11/20) used heuristic based estimations, for example,

*"Um .. well in the first two weeks uh she wasn't really with it . and in the last two weeks, obviously in the last week particularly it was most of the time so I would say often."*

(ID 15; 17b)

## **Depression**

Similar to anxiety, rate based estimations were seldom used in reporting depression frequency. A fair proportion of proxies reported that the patient was depressed "*all of the time*" in the last week (5/13) and month (5/17) of life. Thus, there were few (8/13) reported strategies used to recall the patient's condition in the last week. Of these, half recalled and counted episodes. Again, episodes were recalled for frequent and infrequent episodes. Where episode enumeration was used for persistent depression, like anxiety this tended to be associated with a prolonged episode such as a reaction to an event or observed during visits to see the patient. This was illustrated in one response,

*"Every time you went in there [hospice] she was constantly saying oh I am going to die. So especially after the doctor told her on the Tuesday. So yeah every day most of the time."*

(ID 15; 20b)

Reporting on patients' last month, six of the nine proxies who reported strategies, used heuristic based estimations. For example,

*"It [depression] was now and again .. it changed and became worse over that month but overall it was sometimes"*

(ID 2; 25b)

For pain, anxiety and depression there was evidence of averaging over the reference period and judgments based on the proxies innate sense of frequency. However, without extensive

probing it was not possible to distinguish the actual heuristics used. The findings from proxies reported frequency strategies are consistent with other studies in showing that participants use a variety of strategies rather than just a recall of each episode (Bickart *et al.*, 1994; Blair *et al.*, 1991). Factors such as frequency were associated with a particular strategy but because of the nature of the symptom and the longevity of an episode in some instances proxies recalled specific episodes.

### 5.5.6 Symptom cues and judgments

#### *Pain*

In the last week of life, overt expressions of pain were the most frequently used indicators of pain presence, frequency and severity. These included medications ( $N=34$ ), verbal ( $N=25$ ) and non-verbal behaviour ( $N=29$ ) such as agitation, posture and facial expressions (wincing, grimacing) The code “*medications*” was an important cue to the level of pain both in the week and month before death. This consisted of increases in analgesics, the taking of medication or the need for stronger analgesics. Disruption to the patient’s functioning and normal daily activities was commonly referred to in the week and month reference periods for identifying pain presence and severity. Restricted mobility and reduced participation in their usual activities were observed, from not walking the dog or gardening, to the extent of not wanting to get out of bed. Surprisingly, none of the proxies stated that the patients’ level of consciousness was affected in the month, even though there was more emphasis on non-verbal behaviour ( $N=32$ ) than verbal behaviour ( $N=16$ ). A possible explanation for this finding might be that the patient’s level of consciousness was less affected in the month than in the week, giving the proxies access to more overt behavioural indicators. Alternatively, patients might not have wanted to bother their significant other by talking about pain. Consequently, proxies relied on non-verbal indicators.

Despite the findings in Chapter 4 where control was important in determining levels of pain intensity, little reference was made to it in these investigations. Control was only mentioned in the month and in relation to the presence ( $N=7$ ) and frequency ( $N=5$ ) of pain. One reason for this might be that control is an individual perception experienced by those with the symptom. Since the proxy is an observer, it might be harder for them to infer how effective the control is. This was supported by the finding that control tended to be inferred when the patients ability to cope with the pain was expressed overtly. This applied equally if it was expressed through verbal communication, specific behaviours or medications usage.

Fifteen proxies mentioned the affective components of the pain experience. With two exceptions, these were expressed in patients with significant and severe pain. Manifestation included aggression, anger, frustration, agitation, tearfulness, quietness and distress. These

were seen as a change in the usual character of the patient, as one proxy recalled her husband's behaviour,

*"Then he became aggressive with it, that made me think it must have been severe [pain] for him, for him to be like that because he wasn't normally like that."*

(ID 22; 3b)

Another proxy told how his father had changed as a result of the pain,

*"It [pain] was very distressing because toward the end he started swearing and he never swore.. um .. I didn't think it was my Dad.... He don't swear usually. He was frustrated."*

(ID 24; 8b)

Trait knowledge (knowledge based on the person's characteristics/personality) was important in forming judgments. Knowing what the patient was able to endure and their character was often mentioned when deciding on the presence ( $N=7$ ) and severity ( $N=6$ ) of pain. This is in line with previous findings where participants are asked to report about others (Blair *et al.*, 1991; Menon *et al.*, 1995). Norms were sometimes used *"She didn't get the pain that some cancer patients get"* (ID 30; 1b), but lack of knowledge and experience to base their judgment on norms may have limited this as a basis for judgments. Instead, most proxies anchored their judgments on the presence of pain in the week ( $N=5$ ) and month ( $N=11$ ) on what the patient had been like previously during their illness. In the previous months or weeks.

*"She suffered from escalating pain right the way through... I think it was a bit more controlled in that week because I think in the last week she was really out of it on morphine, but the pain was still there."*

(ID 5; 1b)

Analysis of the severity of pain ratings showed the variable nature of the pain experience during the reference periods. When asked about how proxies decided on the intensity level, proxies used various strategies. Most based their responses on the most severe pain experienced, even if this was infrequent. The most frequently experienced level of intensity followed this, while others were unable to give an explanation for their choice. This highlights the need to be more explicit when asking about symptoms such as pain because of changeability over the reference period.

### **Anxiety**

Seventeen participants reported the presence of anxiety in the week and twenty in the month. Overwhelmingly, references to patients' feelings such as anger, frustration and the need for reassurance were repeatedly used in the week ( $N=28$ ) and month ( $N=30$ ). This was particularly true for the presence and severity of anxiety in the reference periods. One proxy recollected his wife's anxieties.

*“There was a point when she lost the sight of one of her eyes, like, you know. She thought she would lose the sight of the other and her hair began to fall out. And she was always very conscious of her hair, like you know. She was embarrassed and humiliated.”*

(ID 13; 11b)

These feelings were expressed in a number of ways. Verbal communication was the predominant cue to the presence, frequency and severity of anxiety in the week ( $N=25$ ) and month ( $N=21$ ). Few patients spoke directly about being anxious. Instead, anxiety was inferred from their communications, both in its content and frequency. For example, proxies spoke of patients concerns about ensuring matters were in order and that the family was provided for after their death. Others raised issues about suffering with symptoms, in particular pain, loss of functioning, wanting to be at home, losses associated with life events that they would miss, such as children and grandchildren, and uncertainty about the future. The frequency of these communications and the need for reassurance were indicative of anxiety severity and frequency. This is illustrated in the following response,

*“He kept saying there was no hope and saying things like that. We used to say don’t be silly and he used to say I’m dying...I can remember the night before he died, he kept saying he was dying and that was it umm he became very anxious it was obviously playing on his mind.”*

(ID 26; 11b)

In only a minority of instances functioning ( $N=3$ ), medication/help ( $N=5$ ) and coping ( $N=3$ ) were mentioned in the week reference period. This might have been because of the patient’s deterioration in condition and the stage of the illness. In contrast, functioning ( $N=12$ ), medication/help ( $N=12$ ), and control ( $N=9$ ) particularly for anxiety presence and severity were evident in the last month. Functioning was used more in terms of the ability to function in spite of the illness and its consequences. This was closely associated with coping and medication/help, and was apparent at lower levels, infrequent episodes or no anxiety. However, not sleeping and poor appetite were suggestive of anxiety for three proxies.

Non-verbal behaviours and changes in the patient’s demeanour were other cues to their feelings of anxiety in the week ( $N=16$ ) and month ( $N=11$ ). These included “crying”, “jumpiness”, “withdrawing”, “agitation” and “hostility”. As well as small gestures that were significant to the proxies in suggesting that the patient was anxious.

*“I was sat with him one day at the hospital and he grabbed hold of my hand, which wasn’t like him ’cause we’re not that type of family, and he had hold of my hand and he gripped it real tight. Like he was scared. That’s what made me think he was anxious like.”*

(ID 21; 16b)

This example is also illustrative of anchoring the judgment on the patient’s usual character. This type of trait knowledge was used by a considerable number of the proxies in the month

(10/20) to decide whether the patient was anxious or not, and in the week to judge the presence (5/17) and severity (6/17) of anxiety. In the month reference period several proxies anchored their judgments on what the patient had been like previously during their illness in making a decision on the presence (5/20) and severity (5/20) of anxiety. There was less variability in anxiety and depression intensity than for pain intensity. Where variability was mentioned, there was no clear strategy used to reach a decision on the intensity of anxiety and depression. Both the most frequent level and the most intense level were chosen, but many were unable to say why they had decided on a particular response.

### ***Depression***

Thirteen proxies reported depression in the last week and slightly more in the month (17/30). Again references to the patients' feelings in the week ( $N=27$ ) and month ( $N=35$ ) were most often used as indicators for presence, frequency and severity of depression. Amongst these were sadness, anger, flatness of mood and boredom. In contrast to anxiety non-verbal communication ( $N=18$ ) was frequently mentioned in the week of the patient's death. The patient's appearance of unhappiness was a significant cue. Others were "*clinging*", "*lack of interest*", "*inability to concentrate*", "*tearfulness*", "*irritability*", "*tiredness*" and "*withdrawal*". These feelings were seen to affect patients' ability to fully function in usual activities and to disrupt activities of daily living such as sleep patterns. Conversely the disruption to functioning as a result of their illness was a major causal factor mentioned in the patient's depression. Thus functioning was a prominent indication of depression.

While in the month reference period verbal communication was mentioned ( $N=19$ ) slightly more than non-verbal communication ( $N=15$ ). Of the thirteen proxies who verbally communicated with the patient, in four of these cases the proxy said that the patient was not depressed. Only two proxies explicitly talked with the patient about feelings of depression. Similarly, to anxiety most judged the symptom on what the patient talked about. For example, fears about dying and unhappiness with their circumstances. One proxy recalled her mother's depression, "*Well, she kept saying she wanted to die*" (ID 5; 24). Two patients were being treated for depression with antidepressants prior to month.

Again, judgments regarding the presence of depression were often anchored on trait knowledge about the patient's personality and characteristics in the week (11/30) and last month (11/30) reference periods. An actual change in personality was observed in a few instances. For example,

*"Anyway, we managed quite well but the pain became progressively worse and over the last few months she had a lot of pain. I.. well me and my sisters , they both noticed it as well. She changed. Her whole personality just seemed to change".*

(ID 1; 21b)

As was depression severity in the month (5/17). Following this, anchors such as what the patient had been like previously during their illness also signified depression severity and presence in both reference periods. Notably judgments were also based on norms a finding not observed for pain or anxiety. As the following example illustrates,

*“It’s [depression] a natural thing really, because you’ve got cancer”*

(ID 15; 26b)

The focus of the Chapter 6 is to examine proxies emotions and whether they have a bearing on their responses with regard to patients’ pain, anxiety and depression.

### 5.5.7 Event cues

#### *Pain*

Event cues tended to be used to signify the presence of pain, anxiety and depression rather than its intensity and frequency. Places were usually mentioned particularly in the month reference period (10/25). These consisted of references to periods before or after admission or input from a hospital or the hospice. For example, *“I would say that most of the pain was just before she went into hospital”* (ID 13; 8b). This is in line with the memory literature where event cues prompt recall of events (Brown *et al.*, 1986). Several proxies talked of escalating pain and the hospice’s involvement in controlling the patient’s pain. This was not surprising given that uncontrollable pain and difficulty coping were the main reasons for admission to the hospice in the last week of the patient’s life (10/30). Changes of dressing or other treatments were characteristic of events associated with pain. As was the patient’s condition, where pain was seen to intensify nearing death. This is confirmed in the numbers with severe pain in the last week of life (12/30) compared to the month before death (7/30).

#### *Anxiety*

In the week (7/13) and month (2/20) presence of anxiety was associated with a significant event in the lives of the patient and proxy. A major event was admission to the hospice in four cases this was viewed as an extremely stressful experience. As one proxy’s response indicates, when talking about her mother’s transfer to the hospice,

*“She didn’t know what was happening at that stage she was too ill to understand, if she’d gone in earlier we could have explained it to her, it would have been planned instead she was transferred in the night, that terrified her..”*

(ID 30; 11b)

The patient being told their diagnosis and characteristics of events ( $N=11$ ) such as a change and deterioration in the patient’s condition and symptoms, a fall, managing symptoms, dressing changes and the last rites.

## ***Depression***

To a lesser degree than anxiety, major events featured as event cues for depression in the week (5/13) and month (3/17). Being in hospital was an event mentioned as time of depression for three patients. Family gatherings were events where some patients were seen to “*put on a brave face*” to conceal their feelings from other family members. A significant characteristic event associated with depression was deterioration in the patient’s condition and loss of functioning. There was some overlap with anxiety in the cues used. For example, three participants who had mentioned specific events such as a fall, dressing changes and the last rites also used these cues to indicate depression.

### **5.5.8 Reference periods**

There was a tendency for some proxies to report on periods extending the reference periods asked about, for the presence of a symptom. The numbers were relatively small for pain (3/25, week; 2/26, month) and anxiety in the week (3/17). However, for anxiety in the month (5/20) and depression in the week and month several proxies, five and six respectively, spoke of periods from weeks up to several months outside the reference period. This is what Sudman and Bradburn (1973) refer to as *forward telescoping*, which can lead to an increase in estimates of particular occurrences. The clarity or saliency of an event is thought to lead the individual to the belief that the event occurred more recently than it actually did (Brown *et al.*, 1986). Where forward telescoping appeared it coincided with a salient and emotional event such as the patients’ diagnosis or deterioration in the patient’s condition or treatments,

*“I think he had about six lots of treatment. He had loads and loads of treatment and every time .. I think the most anxious or the most when he decided that things weren’t going to go his way was when he was diagnosed as having secondaries.”*

(ID 29; 16b)

In this example, the proxy later adds that this is in the six or seven months before the patient died. This example was typical of the responses where forward telescoping was observed. Most of the proxies who discussed periods beyond the reference periods, spoke of when the symptom started. In two incidences patients had a pre-existing diagnosis of depression. What is not known is whether this information extended beyond the period is used to form a judgment on the frequency and severity of the symptom in the period being assessed. Pain, anxiety and depression are all highly emotional and salient events therefore it was surprising to find little evidence of forward telescoping for anxiety in the last week and pain. A possible explanation might be that the verbal protocols did not reveal all instances where forward telescoping occurred. Unless the proxy explicitly mentioned that the period was outside the week or month, or the response indicated that the proxy was talking about other time periods and this



was verified through follow up questioning, there may have been other instances forward telescoping occurred but were not detected. An alternative explanation might be that pain in the week and month was highly memorable particularly as almost three quarters of the patients suffered with severe or significant pain. Thus, the memorability of events might have resulted in proxies keeping their responses within the parameters.

## 5.6 Conclusion

The findings presented in this chapter suggest that the patient's level of consciousness and ability to communicate create significant difficulties for the proxies who represent them at the end of life. This was most evident in the last week of life and for detecting the presence of depression and anxiety. By comparison, proxies were better able to make judgments regarding the presence of pain as this could be communicated even when the patient was not fully conscious. This was through cues such as moaning, grimacing and agitation. However, the patient's level of consciousness did impede proxies' ability to assess the severity of patients' pain.

The present investigation was exploratory, as there exists no published research to compare the findings with. Examination of the cues proxies use to judge patients' pain, anxiety and depression revealed a wide variety of indicators. These included signs commonly associated with the symptom such as analgesic use, facial expressions and verbalisations of pain; expressions of fear and agitation signifying anxiety; to sadness and a reduction in functioning indicating depression. This suggests that proxies are drawing on different sources of knowledge some of which may be reliable indicators. Although non-verbal behaviour provided important cues to symptoms, proxies also relied to a large extent on verbal behaviour, both in the last week and month of the patient's life. One reason for this might have been the lack of concrete observable phenomenon with the symptoms studied. A notable observation was the use of idiosyncratic expressions familiar to both the patient and proxy. This could be a look or expression, or behaviour out of the ordinary, such as holding hands. Such observations were meaningful to the patient and proxy. However, their indication of an actual symptom is not known. Comparisons between patient and proxies expression and understanding of symptoms would help elucidate this.

Interpretation of proxies' verbalisations revealed that proxies do not rely entirely on overt expressions of symptoms, despite the fact that these often viewed by proxies as the most important indicators. Their judgments are anchored within the context of their general knowledge. This agrees with findings in the cognitive psychology and survey methodology literature (Schwarz, 1990; Blair *et al.*, 1991; Menon *et al.*, 1995; Schwarz & Sudman, 1996). Among the most commonly used general knowledge was trait knowledge based on patients'

characteristics/personality. This guided proxies' decisions on pain, anxiety and depression to a large extent. Of particular importance were changes in the patient's character and behaviour from before their illness. Whilst norms have been observed as a basis for judgments in other surveys (Bradburn *et al.*, 1987), they were seldom referred to in these investigations, probably because of proxies' limited experience and knowledge of terminal illness. Instead, proxies depend on their experiences of what the patient was like previously during their illness to anchor their symptom judgments. This is important because only significant others in close contact with the patient would have access to this information. From this it is possible to surmise that significant others may be better proxies for some aspects of the patients experiences than health professionals who have limited contact with, and information about the patient.

A mixture of strategies were used to judge the frequency of symptoms. In assessing pain and consistent with the survey methodology literature, infrequent episodes of pain were recalled and counted (Blair & Burton, 1987), whereas, for frequent pain episodes the strategy depended on the regularity of the pain. Rate-based strategies tended to be used where the pain showed some degree of regularity, such as at night or four-hourly. Under these circumstances, these estimates may be accurate. In contrast, rate-based strategies were rarely used for depression or anxiety. This might be due to the variable nature of these symptoms. Episode enumeration and heuristics were used almost equally. Recall of episodes coincided with particular events and was recalled with a great amount of detail, suggesting their salience to the proxy. As a result, recall might have been enhanced and accurate.

There was evidence of forward telescoping for all symptoms, especially for depression and anxiety. This might be due to the nature of the symptoms as several proxies spoke of when the symptom began. As Sudman and Bradburn (1973) identify this can result in an increase in estimates. Despite the fact that the reference period was given in each question participants still went outside the period. This highlights the need to clearly define the period and continually remind participants. Several methods have been suggested to overcome this problem although the practicality of some may make them unsuitable within the context of palliative care. For example, bounded recall where participants are repeatedly interviewed and reminded of the events reported would not be feasible (Neter & Waksberg, 1964). A more appropriate method would be to encourage participants to recall events within a personally significant time frame. Here, proxies could report on events such as the patient's experiences following admission to a hospice or hospital, a common event cue used by proxies in this study. This would give variable reference periods, but would be useful for evaluating care within a particular setting. Alternatively, using their personal framework, proxies could first be asked to recall events within context and then asked to place events within the reference

periods. Evidence to support the use of these methods comes from Means *et al.* (1989), where a fifteen percent increase in accuracy of dating doctors' visits was shown.

As an aid to recall, the last month and week of the patient's life were used as reference periods. These periods were chosen based on the cognitive and survey methodology literature. The present research suggests that the period needs to be long enough to assess the phenomenon of interest, but short enough to prevent decay of memory for events. The findings from this study suggest that, because of the variable nature of the symptoms, the reference periods should be kept fairly short. However, the number of difficulties experienced over the week, particular for anxiety and depression, suggest a longer period such as the last month of the patient's life. A recommendation here would be to assess symptoms over a shorter period than the last year of life, which is used in VOICES.

### 5.6.1 Limitations

The limitations of the methods chosen have been discussed in Chapter 3. In this section, the limitations of the study in relation to the aims are discussed. The aim of the study was to investigate what proxies are recalling when they report on patients' levels of pain, anxiety and depression, as means to understanding possible biases and errors. Although the study identified areas where these occurred, such as forward telescoping and heuristic strategies, it is not possible to state from this study that these actually biased proxies' responses. It is only possible to infer their likely effects. In addition to this, relations between the cues used to judge symptoms and the validity of proxies' reports cannot be established, as there is no comparison with the patients whom they are representing.

Proxies varied in their verbalisations of their thought processes. Some proxies were eloquent in their responses, while others experienced difficulties. One reason for this was the amount of contact the proxy had with the patient, reinforcing the need to choose a proxy who has access to the required information. Exploration of the heuristic strategies used by proxies was inhibited by the necessity to avoid distressing proxies. Making these heuristics the subject of investigation in their own right would have been insightful, but this has to be balanced against moral and ethical concerns (discussed in Section 3.5.3). A further limitation was that many of the proxies reported that the patient had depression "*most of the time*"; only eight proxies reported strategies in the last week of the patient's life. Therefore, interpretation needs to be cautious given the numbers.

### 5.6.2 Summary

The present study provided insightful findings on the processes underlying proxies' judgments of patients' pain, anxiety and depression. It also raises awareness of the complexities of

the mental processes used by proxies and the importance of social processes in human decision-making. Furthermore, within the context of studies comparing patients and proxies, it helps elucidate possible reasons for biases. Of importance for the retrospective approach it shows that proxies amass information about patients symptoms from a variety of sources questioning Higginson *et al.*'s (1994) claim that proxies accounts are based on their own feelings. However, further research is needed before this can be disputed. Hence the next chapter examines proxies' emotions in relation to their responses. Throughout the present investigations, areas for future study have been mentioned and recommendations are made to improve the design of the VOICES questionnaire. However, the study presents researchers using the retrospective approach with practical points that are useful at the design stage of survey development.

# CHAPTER 6: Proxies' emotions and judgments

## 6.1 Introduction

The death of a loved one evokes a variety of feelings, emotions and thoughts, which are characterised as grief. Since the retrospective approach relies on collecting information during this process, it has been asserted that the proxies' responses may be a reflection of their own feelings rather than the experiences of the patient (Higginson *et al.*, 1994). To date, there has been little attention given to the effects of bereavement on proxies' responses. As noted in Chapter 2, this may be because of the difficulties in conducting this type of research. Only one study has investigated proxies' emotions in relation to their retrospective responses. Here Cartwright and Seale (1990) analysed data from Cartwright *et al.*'s (1973) study to examine whether proxies' experiencing anger responded differently from those who were not. Information regarding the patient's death, such as cause of death, age and place of death, symptoms, care and help received were collected. No significant differences were identified. Given the lack of empirical evidence, the aim of the present investigation is to explore whether proxies own emotions are reflected in their reports of patients' pain, anxiety and depression. In addition to this, caregiver strain is included as a number of studies, highlighted in Chapters 1 and 2, suggests that the level of strain is related to the level of congruency between patients and proxies' reports of patients' symptoms. The proposed investigations may shed light on this relationship. Furthermore, the negative consequences of caring include emotional distress and psychiatric morbidity (Schulz *et al.*, 1990; Siegal *et al.*, 1991). Relevant background literature drawn from fields such as palliative care, cognitive, social and clinical psychology is discussed in relation to this issue.

### 6.1.1 Grief

Grief is a natural and complex multidimensional response to loss that can continue for several years. However, most researchers agree that typically there is improvement within the first one to two years (Stroebe, *et al.*, 1993, for discussion). Multiple variables have been found to influence the manifestation of grief. These include the timing of the loss, degree of attachment, relationship with the deceased, cultural background, personality, social network, gender, religiosity and socioeconomic status (Parkes, 1975; Rubin, 1985; Parkes, 1988; Vachon & Stylianos, 1988; Rosenblatt, 1993; Stroebe & Stroebe, 1993). The process of grieving has been characterised in various ways depending on the theoretical position of the researchers (see Shakleton, 1984; Hogan *et al.*, 1996 for reviews). However, two stage theories have predominated; Bowlby's (1969) attachment theory where grief has been conceptualised as a general response to separation in an attachment bond (Parkes, 1965; Bowlby & Parkes, 1970;

Bowlby 1980), and Kubler-Ross' (1969) stages of dying that has been applied to grief. Both theories conceptualise grief responses in a series of stages or phases but differ on the number and content of some of the stages. For instance Bowlby (1980) put forward four phases, numbness/disbelief (distress and/or anger), protest (anger and anxiety), despair (depression), and acceptance. Whereas, denial, anger, bargaining, depression and acceptance form Kubler-Ross' (1969) five stage account. Although the stage theories of grief have been valuable in informing experiences during bereavement, they have criticised for being static and linear (Wortman & Silver, 1992; Shuchter & Zisook, 1993). Furthermore, empirical evidence does not always support the stages proposed (Silver & Wortman, 1980). Instead, after the initial shock of the loss, the process of grief is viewed as much more changeable with a mixture of responses which vary from individual to individual (Shuchter and Zisook, 1993).

### ***Anxiety and depression in grief***

Anxiety and depressive symptomology are common in bereavement. These can occur in the recently bereaved and last for varying lengths after the patient's death (Parkes, 1965). Compared to non-bereaved controls anxiety symptoms have been found to be higher during the first six to seven months of bereavement (Bartrop *et al.*, 1992; Zisook, *et al.*, 1990). Similarly, depressive symptoms have been shown to be higher in the month after death, after which levels decline over six to twelve months (Parkes, 1965; Vachon *et al.*, 1982; Parkes & Weiss, 1983). However, other researchers have identified elevated levels of anxiety and depressive symptoms approximately a year following the death of a significant other (Zisook *et al.*, 1990; Zisook & Schuchter, 1991). Even though anxiety and depression symptoms are prevalent, they are not always at clinical levels. In a study of 350 bereaved spouses, responses characteristic of anxiety and depression, such as fearfulness, crying, apathy, difficulty concentrating and feeling overwhelmed, were frequently endorsed by participants (Shuchter & Zisook, 1993). Due to the overlap with these types of responses and depression it can be difficult to differentiate depression from grief reactions. Nonetheless, a substantial number of bereaved individuals' symptoms reach clinical levels for the diagnosis of a psychiatric disorder (Jacobs *et al.*, 1989; Jacobs *et al.*, 1990; Jacobs & Kim, 1990). Jacobs *et al.* (1990) found that, in the first year of bereavement, 30% met the criteria for generalised anxiety disorders and 10% for panic disorders. Depression levels were as high as 32% in the first six months, and 27% at 12 months post bereavement for spouses (Jacobs *et al.*, 1989). At these intense levels, grief is likely to be considered atypical (StroebeStroebe *et al.*, 1993).

### 6.1.2 Cognition and emotions

Having discussed grief, the focus now turns to how emotions can influence the way individual's process information about events. From the psychological literature there are differing theories about the relationship between cognition and emotion, which centre on whether they are part of the same system or independent systems (Bower, 1981; Wyer *et al.*, 1999). However, there is consensus based on the empirical evidence that the relationship is bi-directional and complex (Wyer, *et al.*, 1999, for discussion). One conceptualisation of the relationship is that emotions themselves are elicited and linked to the evaluation of goals and their outcomes (Lazarus, 1990; 1991; Stein & Levine, 1990). As Stein and Trabasso (1992) point out, anger may be expressed when there is failure to attain a goal and the focus is on the cause, whereas sadness may result from focusing on the consequences of goal failure. These evaluations occur within a social context, and as such, are likely to be influenced by the individual's culture, expectations and beliefs. There are also strong links between emotion and information processing, where memory and judgments have influenced the encoding, retrieval and evaluation of information (Bower, 1981; Beck *et al.*, 1985; Schwarz & Clore, 1988). Since the evidence is pertinent to the present investigations, relevant aspects will be discussed in more detail.

#### *Emotions and memory*

Emotions are an integral part of attention, perception and retrieval of information. As discussed in Section 2.4.2, attention is drawn to central aspects of highly emotional events, which remain poignant for several months after the event (Christianson & Loftus, 1987; Christianson, 1992). Evidence from non-clinical and clinical populations also suggests that the emotional state of the individual can bias attention to stimuli (information, events) consistent with their mood-state (Beck, *et al.*, 1985; Broadbent & Broadbent, 1988; Burke & Matthews, 1992). Important links between emotion and cognition are proposed in Bower's network theory (1981; 1987), which is used to explain and predict mood congruency. Blaney (1986) defines this as "*enhanced coding and/or retrieval of material the affective valence of which is congruent with the ongoing mood*" (p.229). Within Bower's (1981; 1987) theory, emotions are part of the retrieval of events increasing accessibility to memories of a similar affective tone. There is considerable evidence to support mood congruency effects (e.g. Bower, 1981; Gilligan & Bower, 1984; Watkins *et al.*, 1992). However, the findings are sometimes inconsistent and have been shown to vary depending on the conditions (Blaney, 1986; Parrott & Sabini, 1990).

## *Emotions and judgments*

In Chapter 5, the processes underlying judgments were discussed. For example, the basis for judgments (e.g. normative behaviour and general knowledge), cues and estimations strategies (e.g. heuristics, enumeration) were explained. Since the basis for judgments are embedded in individuals' general knowledge about a particular event, they may be influenced by the proxies' own emotions. This may be particularly evident when other criteria are not available or difficult to apply, for example when the information cannot be retrieved from memory or the judgments are difficult or ambiguous (Schwarz & Clore, 1988), and in depressed individuals (Williams & Dritschel, 1988). Under these circumstances, individuals may rely on heuristics (e.g. rule-of-thumb methods). One of which is the *how-do-I-feel-about-it?* heuristic proposed by Schwarz and Clore (1988). Here emotions are used as a source of information for making judgments about the self, other people, objects or events (Forgas, 1991). These affect-based evaluations, as Wyer *et al.* (1999) points out are almost exclusively based on the individual's feelings about events. These contrast with judgments based on general knowledge where evaluations are based on knowledge about the person, symptoms being judged and observations. There is some evidence to support this heuristic within the research into self-report, where assessment of mood states and pain are biased toward present beliefs (Bryant, 1993).

Cognitive theories of emotional disorders have highlighted the distortion of perception, memory and appraisal in patients with anxiety and depression (Beck *et al.*, 1979; 1985). Here, perception of new information is organised within a particular internal representation (schemata). As Williams *et al.* (1997) points out, the key schemata in patients with anxiety are threat and vulnerability, whereas negativity is prominent in the schemata of patients with depression. Correction of these biases or faulty thinking has been the impetus of much of the cognitive therapy techniques used to treat patients (Beck, 1963; 1979; Beck, *et al.*, 1985). Despite the fact that much of the work has been conducted outside the field of palliative research, the findings provide valuable insights and may have implications for the use of proxies in retrospective surveys during bereavement.

## **6.2 Aims**

Proxies' responses to the PSCI and ratings of patients' pain, anxiety and depression will be interpreted in the light of their self-reported depression, anxiety and caregiver strain. This will help to establish whether proxies own emotions are reflected in their responses or are used as a basis for judgments.



*Research question three*

Do proxies’ own emotions have a bearing on their reports of patients’ pain, anxiety and depression?

**6.3 Methods**

The methods and procedures are detailed in Chapter 3 and Section 5.4. This involved analysis of proxies’ responses to the PSCI. The codes used are shown in Table 6.1 and are based on the literature regarding affect-based judgments (Schwarz & Clore, 1988; 1989; Wyer *et al.*, 1999). Affect-based judgments were instances where proxies’ responses were based on their feelings about events or the patient, rather than other indicators such as observations, communication and behaviours. In addition to this, proxies’ verbalisation of their emotional responses were also coded. Criterion sampling was used based on proxies’ scores on the Caregiver Strain Index (CSI) (Robinson, 1983), State Anxiety Inventory (SAI) (Spielberger *et al.*, 1983) and Beck Depression Inventory (BDI-II) (Beck *et al.*, 1996). The sampling criterion will be discussed along with the analysis in Section 6.4.2.

**Table 6.1: Content analysis codes for proxies’ emotional reactions and affect-based judgments.**

|  |
|--|
| <i>Proxies’ emotions</i>   |
| Proxy’s emotional reactions to event mentioned   |
| <i>Affect-based judgments</i>  |
| Explicit inferences about how the patient was feeling or thinking based on proxies own feelings about events |

**6.4 Results**

**6.4.1 Descriptive statistics**

Tables 4.1 and 4.2 contain summaries of patient and proxy characteristics and demographics.

*Proxy self-report measures*

Scores on the CSI were out of a possible maximum of thirteen. Half of the sample scored  $\geq 7$  indicating moderate to high levels of perceived caregiver strain. The mean score was 6.7 (SD 2.4), with a range of 3-12. Self-reported anxiety using the State Anxiety Inventory (SAI) was above average (mean 41, SD 12.4) compared with adult populations of males (mean 35.72, SD 10.4) and females (mean 35.20, SD 10.61). Scores ranged from 22-69. Thirty percent (9/30) of proxies scored higher than the normative range for males (46.21) and females (45.81). Proxies’ reports of their levels of depression using the Beck Depression Inventory (BDI) ranged from 0-31 (mean 10.1, SD 6.8). The majority (26/30) had scores showing normal or mild levels of depression. Two proxies’ scores indicated mild to moderate depression (BDI

scores 16-19), one proxy scored twenty-seven indicating moderate to severe depression and only one proxy had a score (31) denoting severe depression.

### ***Bivariate correlations***

As expected, there was a significantly high correlation ( $r=.82, p<.01$ ) between scores on the SAI and BDI. A relationship was observed between the proxy's age and their scores on the SAI and BDI, with younger proxies scoring higher on the SAI ( $r=-.57, p<.05$ ) and the BDI ( $r=.37, p<.05$ ). Proxies' CSI ratings were found to be unrelated to SAI or BDI ratings. In addition, they were not associated with any of the patient or proxy demographics or characteristics.

## **6.4.2 Analysis procedure**

Content analysis of the PSCI and the reliability and validity of the coding are reported in Sections 5.5.2 and 5.5.3. Qualitative description of the content analysis is reported below. Quantitative analysis of the data involved a comparison of proxies' self-reported caregiver strain, anxiety and depression with; (i) the frequencies of the codes (proxies' emotions and affect-based judgments) in their responses, (ii) proxies ratings of patients' pain, anxiety and depression using VOICES. This analysis was conducted to investigate whether there was a relationship between their responses/ratings and emotions.

### ***Qualitative description of the content analysis***

***Proxies' emotions*** incorporated those instances where the proxy made reference to their own feelings within their responses when they were asked about the patient's symptoms. For example, when asked about depression one proxy stated, "*I would. I mean I was getting upset*" (ID20; 19d). Proxies' own feelings were more evident when asked about patients' anxiety as compared to pain and depression. A variety of emotions were expressed with many conveying empathy for their significant other and their suffering. The condition of the patient caused the proxy considerable distress, as the following quote illustrates.

*"It was really upsetting for me and my sister seeing him in that state, we like to remember him as he was."*

(ID21; 11b)

Sadness was also expressed at the loss of the patient's control, functioning and future life. Attributions of blame and responsibility were focused on the self and the health care system. Pertaining to the self were feelings of guilt, helplessness and regret at actions or failures to act. One proxy spoke about how guilty she felt for not believing that her husband was dying and for not taking the matter more seriously (ID 26). There was also regret and guilt for not being more forceful with health carers and for their sense of helplessness at not knowing how

best to proceed. The daughter of one patient spoke about how she had withheld doses of opioid analgesics from her mother, for fear of overdosing her. This had resulted in her mother being in considerable pain (ID 5). In several instances, proxies vocalised their anger at health professionals involved with the care of the patient. Anger resulted from not being listened to, loss of control, lack of information and the standard of patient care. Three proxies talked about their own distress when the nurses came to change dressings. Shock and disbelief were evident in two cases where the death was sudden and the patient was young. Although most of the reactions were negative and throughout many of the interviews proxies were tearful, some reactions were more positive. Amongst these were satisfaction with care, humour, relief that events were over, and a positive outlook for the future.

*Affect-based judgments* included instances where proxies' responses were based on their feelings about events or the patient, rather than other indicators such as observations, communication and behaviours. Since pain, anxiety and depression are considered more subjective in nature, affect-based judgments were common, as the proxy attempted to view events from the patient's perspective. The son of a patient with a hearing and visual impairment whose functioning had declined, spoke of his father's depression,

*"Thinking of him sat there in that house, sat in the chair, waiting and it was such a silent world and that distressed him most. Just sitting there, hours and hours, between when his carer left to when the carer came back. I can't imagine how awful it must be. It must be terrible."*

(ID 24; 24b)

Many of the relationships were long standing and as a consequence knowledge of symptoms was based on similarities between how the proxy felt and how the patient felt or on the intuition gained from their life together and their shared experiences. Further prompting did not reveal the basis for the intuition other than it was a feeling that the proxy had about the patient's experiences. As one proxy responded, *"Um [sighs] it's hard to explain really. Just a feeling about my mother"* (ID 5; 11b).

When asked about the patient's pain, the husband of another patient replied,

*"Um having been married thirty years I knew she was in pain. You get to know someone that you live with and you know when they're in pain."*

(ID25; 1c)

The words *"us"* and *"we"* were also indicative of shared feelings. The bond between the patient and the proxy meant that proxies felt able to make judgments based on their own assumptions. This was apparent in some instances where the proxy stated that the patient was unaware of their condition due to their level of consciousness, knowledge about prognosis or lack of communication about terminal illness. Since the patient was not aware, anxiety and depression were not considered a problem. In contrast, awareness of circumstances led some

proxies to believe that the patient was in pain, anxious or depressed, as they would have been given the same circumstances. Proxies’ own concerns were apparent in a small number of cases. This was exemplified in the response of a husband talking about his wife’s pain,

*“Well, I would put it very, very high. That was my main concern. When I knew that things were really hopeless, I was concerned that when she had it [pain], I could almost experience the pain myself.”*

(ID 13; 8b)

Another instance was a proxy talking about her mother’s depression,

*“I would say it was significant ..um. because .. like anyone you would be wouldn’t you? If you, if you don’t know what is happening and you see lots of different people traipsing in and out, lots of different district nurses and all the rest of it.”*

(ID 5; 26b)

***Quantitative analysis of frequency of codes and self-report measures***

In order to examine whether there were any differences between those with high and low CSI, SAI and BDI scores in relation to the proportions who invoked the codes and those who did not, a Fisher’s exact test (two-tailed;  $\alpha=.05$ ) was performed. Groups were divided based on the CSI and SAI into high and low groups signifying the level of caregiver strain and anxiety. The numbers reporting mild or no depression ( $N=26/30$ ) from the BDI scores meant that there were too few in the high group to perform analyses. Scores on the CSI of  $\leq 6$  and  $\geq 7$  were used to divide the sample into low and high groups respectively. Division of the sample for the SAI was based on the population averages for males and females as detailed above. In all but two instances, the proxies with higher levels of caregiver strain and anxiety were more likely to invoke both codes. The only exceptions were where proxies with low CSI scores invoked the proxies’ feelings code for pain and depression more often compared to those with high CSI scores. However, none of the differences were statistically significant.

The absolute frequency or number of times the code was applied was also compared between proxies with high and low scores. Proxies in the high CSI and STAI group tended to apply both codes more often, with the exception of depression in the low CSI group (Tables

**Table 6.2: The absolute frequencies of codes for CSI.**

| Code       | Affect-based judgments |                    | Proxies’ feelings   |                    |
|------------|------------------------|--------------------|---------------------|--------------------|
|            | High CSI ( $N=15$ )    | Low CSI ( $N=15$ ) | High CSI ( $N=15$ ) | Low CSI ( $N=15$ ) |
| Pain       | 20                     | 13                 | 7                   | 4                  |
| Anxiety    | 25                     | 24                 | 15                  | 9                  |
| Depression | 24                     | 17                 | 5                   | 10                 |

**Table 6.3: The absolute frequencies of codes for SAI.**

| Code       | Affect-based judgments |                | Proxies' feelings |                |
|------------|------------------------|----------------|-------------------|----------------|
|            | High SAI (N=18)        | Low SAI (N=12) | High SAI (N=18)   | Low SAI (N=12) |
| Pain       | 22                     | 11             | 9                 | 2              |
| Anxiety    | 34                     | 15             | 13                | 10             |
| Depression | 28                     | 13             | 11                | 4              |

6.2 & 6.3). Analysis of the frequencies using the Mann-Whitney test (two-tailed;  $\alpha = .05$ ) indicated that none of the differences were statistically significant.

***Quantitative analysis of VOICES rating and self-report measures***

In order to statistically analyse the VOICES ratings in relation to proxies' self-reported levels of caregiver strain and anxiety, the ordinal ratings from VOICES were assigned a score such that on the frequency scale 5 = *most of the time* to 1 = *never*, and on the severity scale 5 = *severe* to 1 = *none*. As the data violated the assumptions of parametric analysis, the Mann-Whitney test was used. In all instances, the Kolmogorov-Smirnov (K-S) two-sample test ( $\alpha = .05$ ) showed that the data met the assumption of the Mann-Whitney test for similarity of distributions. A series of two-tailed Mann-Whitney tests were performed. Due to the number ( $N=24$ ), only those where a significant difference was found, are reported (full analysis contained in Appendix O). In nearly all the analyses, proxies in the high CSI and SAI groups rated patients' symptoms as more severe and intense and several of these almost reached significance. Examples for the CSI ratings existed for pain severity and frequency in the last week and month of life. However, only pain frequency in the month was significantly different ( $z = -2.20, p = .03$ ; K-S  $z = 1.10, p = .18$ ), with the mean ranks for CSI high (18.90) and low

**Table 6.4: Analysis of differences in VOICES ratings between those with high and low levels of self-reported anxiety.**

|                             | SAI High |           | SAI Low |           | Mann-Whitney |     | (K-S) |
|-----------------------------|----------|-----------|---------|-----------|--------------|-----|-------|
|                             | N        | Mean Rank | N       | Mean Rank | z            | p   | p     |
| Anxiety frequency (week)    | 18       | 18.64     | 12      | 10.79     | -2.48        | .02 | .08   |
| Anxiety severity (week)     | 18       | 18.25     | 12      | 11.38     | -2.18        | .04 | .08   |
| Anxiety frequency (month)   | 18       | 18.31     | 12      | 11.29     | -2.22        | .03 | .23   |
| Depression frequency (week) | 18       | 18.06     | 12      | 11.07     | -2.07        | .05 | .12   |
| Depression severity (week)  | 18       | 18.28     | 12      | 11.33     | -2.22        | .04 | .12   |

Two-tailed,  $\alpha = .05$ .

groups (12.10). Statistical analysis comparing proxies in the high and low SAI groups and their VOICES ratings revealed that they were different in several aspects. As Table 6.4 shows, proxies with higher levels of self-reported anxiety rated patients' anxiety and depression in the week, and anxiety frequency in the month as more frequent and severe than those with lower levels of self-reported anxiety. In addition to this, anxiety severity in the month almost reached significance ( $p=.08$ , *n/s*).

## 6.5 Discussion

In the sample of proxies' studied, self-reported caregiver strain and anxiety were high. Given that most of the proxies were the primary caregiver before the patient's death or admission to a hospice, this is not unexpected. The strain of caregiving can impose significant burdens and can cause emotional distress before and after the patient's death (McCorkle, 1988). Moreover, anxiety is commonly experienced during the grieving process, as highlighted earlier (Parkes, 1965; Bartrop *et al.*, 1992; Zisook, *et al.*, 1990). Surprisingly, few proxies' scores on the Beck Depression Inventory were above the cutoff points indicating normal or mild depression, even though elevated levels of depressive symptomology frequently occur during bereavement (Parkes, 1965; Vachon *et al.*, 1982; Parkes & Weiss, 1983). Consequently, the investigations focused on caregiver strain and anxiety.

By dividing the sample into groups based on their CSI and SAI scores, it is clear that there are significant differences between proxies based on their levels of caregiver strain and anxiety and their ratings of patients' symptoms using VOICES. In almost all instances, proxies in the high CSI and SAI groups rated patients' symptoms as more severe and frequent when compared to those in the low CSI and SAI groups. Caregiver strain appeared to have an effect on pain ratings because the differences either reached or almost reached significance (even with the small sample size). Difficulties experienced by the proxy in managing patients' pain may have been a contributor to the results found here. In several of these cases, the patients had to be admitted to the hospice because of uncontrollable pain and other symptoms. Physical care and coping with symptoms can be one of the most demanding aspects of caring for terminally ill cancer patients (Holing, 1986), and is frequently given as a reason for referral to inpatient palliative care services (Seamark *et al.*, 1996). With regards to anxiety, it was noted that proxies' levels were related to their ratings of patients' anxiety and also to depression in the last week of life. A possible explanation for this linkage might be that the patient's anxiety and depression caused distress to the proxy, thus leading to their own elevated levels of anxiety. It is known that patients' distress and symptoms can adversely affect their caregiver's psychological adjustment (Sales, *et al.*, 1992). Alternatively, another explanation that

supports Higginson *et al.*'s (1994) assertion is that proxies' responses may be a reflection of their own feelings rather than the experiences of the patient.

To gain a better understanding of the processes underlying proxies' responses, and to clarify the relationship between proxies' levels of caregiver strain and anxiety in relation to their ratings of patients' symptoms, a qualitative analysis of their responses to the PSCI was conducted. This analysis showed that there was a general tendency for those with greater caregiver strain and anxiety to mention their feelings and use affect-based judgments. However, this did not reach statistical significance and may have arisen by chance. Empathy was frequently expressed as proxies attempted to view events from the patient's perspective. Understandably, this caused proxies distress and sadness. Loss of control and hopelessness at the situation and in dealing with the patient's pain were frequently mentioned.

The present findings may be valuable in explaining the results from several studies that have found that proxies' intensity of caregiver burden is related to incongruency between patient and proxy ratings of patients' symptoms (Kurtz *et al.*, 1996; Miaskowski *et al.*, 1997; Sneeuw *et al.*, 1998). Since proxies' inability to deal with the increasing demands of caregiving creates a variety of feelings, such as guilt and helplessness that are apparent in their thinking when judging the patient's pain. In some instances, this could lead to an overestimation of patients' symptoms, as seen in the studies cited.

Although it is not possible to discern the influence of proxies' feelings on their responses, it is clear from this investigation that their own feelings formed part of their thinking when judging patients' pain, anxiety and depression. This was supported by the identification of affect-based judgments in their responses, where their own concerns and feelings were used as a basis to judge patients' symptoms. An example was knowledge based on similarities between patients and proxies, or an intuitive sense of the patient's experiences in long-standing relationships. In some cases, similarities may provide an accurate representation of patients' experiences, but this cannot be assumed. Furthermore, as the situation is novel to both, the patient may not respond as expected and this may lead the proxy to make erroneous judgments. In light of the findings that there is a propensity for those with greater caregiver strain and anxiety to express their feelings and use affect-based judgments, there is a necessity to take this into account when using information derived from proxies. Considering the findings, the influence of proxies' anxiety on patients' symptoms is clearly an area that requires further investigation. Moreover, as anxiety and depression were significantly correlated in this study, the influence of depression also needs to be explored.

Theoretically the results can be explained within the psychological literature in a number of ways. The relationship between proxies' anxiety and patients' anxiety and depression could be attributed to attention and mood congruency biases in memory (Christianson,

1992; Bower, 1981; 1987). Similarly, cognitive theories of emotion, such as Beck's schema theory (1979; 1985) may explain why proxies' mood states may influence their interpretation of events. Examining proxies' responses there was strong evidence to support Schwarz and Clore's (1988) *how-do-I-feel-about-it?* heuristic or affect-based judgments, which center on how the proxy feels about events. A factor probably contributing to the use of affect-based judgments was the subjective nature of the symptoms investigated and the difficulty of the task. Given that emotions are variable over time, current appraisals of events in a specific mood may change as a function of the stability of the emotion. In Chapter 7, the consistency of proxies' responses over time will be explored.

### **6.5.1 Limitations**

There are several limitations to the research conducted. Firstly, the investigations centred on verbal expression of emotion through the retrospective verbal protocols. A possible explanation for the findings is that proxies, who are more verbally expressive and candid about their own feelings with regard to caregiver burden and anxiety, might be more likely to express their own feelings in their responses to questions about the patient's symptoms. It is also important to highlight that emotions can be expressed non-verbally through behaviours and facial expressions, and physiologically. These manifestations may not be associated with one another. Evidence to support this comes from verbal-autonomic dissociation observed in studies where subjective distress is rated as low, but physiological arousal is elevated when talking about loss (Newton & Contrada, 1992; Bonanno *et al.*, 1995). Indifference may be present in the severely depressed (Watts, 1992). Thus, emotions may be present but not revealed in the present study as the emphasis was on verbalisation. Added to this, repressive coping styles may limit expression and thinking about aspects surrounding the patient's death. Secondly, proxies may not be consciously aware of the effects of some feelings, even though they have been shown to affect judgments (Winkielman, *et al.*, 1997). Therefore, the effects of some emotions may not be amenable to investigation using the verbal protocol approach. Thirdly, caregiver strain is not itself an emotion it is associated with a number of emotions; in these investigations it was not related to proxies' anxiety or depression but to expression of feelings such as helplessness and affect-based judgments. What aspects of the perceived burden of caring influence proxies' responses are not known.

### **6.5.2 Summary**

Throughout the thesis the importance of making accurate judgments about patients' symptoms has been highlighted. It is not possible to say with any certainty whether the proxies' feelings influence their judgments from these investigations. However, identification of prox-



ies' feelings and judgments based on proxies' feelings suggests that they may influence their perception of the patients' experiences at higher levels. The next chapter will consider the consistency of proxies' responses over time.

# CHAPTER 7: Proxies' responses over time

## 7.1 Introduction

There has been very little proxy research wherein the effect of elapsed time has been considered and/or specifically studied. Chapter 2 cites several exceptions. For example, Sneeuw *et al.* (1997a; 1997b; 1998; 1999) in their prospective studies of proxies employed repeated measures designs to investigate the responsiveness of proxies to patients' quality of life over time. Others have investigated changes in the proxy's responses both before and after the patient's death (Higginson *et al.* 1994; Hinton, 1996), or have retrospectively compared proxies interviewed at three months with those interviewed at nine months (Cartwright *et al.*, 1973). From these studies, changes have been shown to occur with the passage of time. For prospective studies, some change is to be expected, as it is likely related to the patient's condition. However, when using proxies to evaluate care retrospectively, changes in their responses could seriously undermine their validity. These issues will be discussed and the effects of time on proxies' responses during bereavement will be explored.

### 7.1.1 Changes over time

Of the little evidence that does exist, there are two studies that indicate that proxies' responses do change with time (Higginson *et al.* 1994; Hinton, 1996). As discussed in Chapter 2, proxies' evaluation of patients' mood showed that depression and anxiety ratings were lower than their prospective ratings, and in some cases, anxiety ratings were closer to patients' ratings (Higginson *et al.* 1994; Hinton, 1996). In contrast, pain ratings became more severe retrospectively when compared to proxies' prospective ratings and to those of the patients themselves (Hinton, 1996). In some cases, the ratings were even polarised to the extremes of the rating scale (Higginson *et al.*, 1994). These studies provide some evidence to support the notion that proxy ratings can change during the terminal and bereavement phases. What is not known is whether and how proxies' accounts change during bereavement when retrospective end of life studies are conducted. To date, only one study, conducted by Cartwright *et al.* (1973), has included a comparison of proxy interviews performed at different time periods during bereavement. The aim of the study was to understand how the timing of interviews affected recruitment, accuracy of information, certainty in responses and feelings about the place of death. It consisted of comparing information gathered at three months and nine months after death. Unfortunately, different groups were compared so it was not possible to assess how proxies' responses may have changed with the passage of time.

### 7.1.2 Timing of interviews

At present, almost all retrospective studies have been conducted with few reasons given for the timing following the patient's death. Other than the obvious difficulty in obtaining tightly controlled and representative samples, the only other comment investigators often make on the issue of timing is the ethics of interviewing proxies too soon. On this latter issue, several studies have used three months into bereavement as a guideline to contact relatives for research purposes (Hinton, 1996; Cartwright & Seale, 1990). Cartwright and Seale (1990) base this timing on Parkes' (1972) study that showed that tearfulness in widows declines substantially between one and three months. However, interviews have been conducted at varying lengths following the patient's death, from as little as eight weeks for information pertaining to patient care (Addington-Hall *et al.*, 1991) up to several years for information regarding morbidity and medical history (Tepper *et al.*, 1993; Rogot & Reid 1975). Notwithstanding, most retrospective palliative care surveys are conducted on average between 3-12 months after the patient's death (Cartwright *et al.*, 1973; Cartwright & Seale, 1990; Addington-Hall *et al.*, 1995a; Addington-Hall & McCarthy, 1995c; McCarthy *et al.*, 1997; Addington-Hall *et al.*, 1998; Teno *et al.*, 2001).

Cartwright *et al.*'s (1973) study contains several findings that are pertinent to the issue of timing. In particular, they found that more refusals as a proportion of failures (62%) were evident at 3 months after death compared to 9 months (48%), although a greater proportion of people could not be contacted at 9 months because no one was living in the area. The importance of memory on the accuracy of recollections and how it affects timing is another topic that has not received much attention within this context. In Cartwright *et al.*'s (1973) study, factual information regarding the patient's death, such as cause of death, age and place of death were as accurate on both occasions, regardless of the timing. Similarly, no significant differences were found for symptoms, care or help received. They also assessed adequacy or certainty of responses, as it was thought that this might be affected. Overall, the findings did not support this, but a smaller number of proxies at 3 months were uncertain about the place of death. Since few details were given for either study, the methodological rigour and analysis could not be critically evaluated. Based on the authors' own conclusions, only small differences were apparent due to the timing of the interview. Considering this evidence, and that from Higginson *et al.* (1994) and Hinton's (1996) studies, a more complete exploration of whether and where discrepancies occur is required. This would be best achieved interviewing the same proxies at different periods during bereavement. This would be very useful for identifying areas where accounts differ and the possible reasons why, and would be instructive for those conducting retrospective studies that rely on proxies. Before exploring these issues a

brief overview of the relevant literature is presented. Within this are factors that could influence individuals' responses, thus causing them to change.

### 7.1.3 Factors likely to affect proxies' responses over time

#### *Memory*

Retrospective reports by proxies rely heavily on their memory for events. These memories are autobiographical in the sense that the proxy, through active participation in the care of the patient or through observing events experiences them. Furthermore, research that has examined memory for self and memory for others suggests that they are fundamentally similar (Kihlstrom *et al.*, 1988). Therefore, in order to understand some of the errors in retrospective reports, it is first necessary to appreciate the significance of autobiographical memory. In previous chapters, errors in survey reports were discussed as they related to retrieval strategies, reference periods and factors associated with the event itself, such as saliency and frequency. In relation to time, it is clear from the literature that memory for events declines over time (Baddley, 1990). Investigations into the subjective qualities of memories, such as perceived completeness and vividness, also support a decline directly following an event, with a more gradual decline after weeks (Thompson, 1982; Thompson *et al.*, 1996) and months (Friedman & deWinstanley, 1998), with little further decline between three and six months (Friedman & deWinstanley, 1998). It should be noted that the personal events studied in the investigations mentioned (Thompson, 1982; Thompson *et al.*, 1996; Friedman & deWinstanley, 1998), involved thanksgiving celebrations and experiences with college roommates, which are far less momentous than the events surrounding the death of a significant other. It may be that the significance of the event along with rehearsal of events through thinking or talking may increase memorability for events increasing retention. Alternatively, coping strategies such as distraction or not wanting to think about events may reduce the intensity of the events and their distinctiveness in memory.

In the aforementioned studies, the focus was on the consistency of reports over time. The literature from both self-report (Schrader *et al.*, 1990; Bryant, 1993) and proxy report (Higginson *et al.*, 1994; Hinton, 1996) suggests that symptom ratings of pain, anxiety and depression can change over time. These findings are consistent with the psychological literature where memory for events is thought to change over time (Conway, 1992; Brewer, 1986; 1996; Thompson *et al.*, 1982). The way in which autobiographical memory is structured explains these changes. For instance, Conway's model (1992) posits that events are stored in hierarchical knowledge structures containing different types of information rather than as episodes. At the top level the knowledge base are plans, themes and goals, within this are general knowledge, knowledge about self and significant others over periods of up to years. The mid-

dle level contains knowledge about specific events over shorter periods, such as months or weeks, which can be thematically organised. At the bottom of the hierarchy, specific events are stored as sensory/perceptual knowledge (thoughts, images and sensations) that is associated with a specific event. With each retrieval, memory is reconstructed by complex processes, each of which are influenced by situational demands (Conway, 1997). This accounts for differences in memories and for the differences in the types of knowledge retrieved. As the analysis in Chapter 5 illustrates, the recall of events included great detail and proxies' thoughts about a specific event, and decontextualised information; normative and trait knowledge. These constructions may play a functional role in adapting to a situation. For example, Taylor's (1991) mobilisation-minimisation theory may explain the changes that result in a decrease in ratings over time. Here, adaptation processes seek to minimise the intensity of negative events such as loss.

### ***Judgments***

References to the inextricable relationship between memory and judgments were made in previous chapters. It is clear from the literature that the fallibility of memory can lead to bias in judgments (e.g. availability heuristic, mood-congruence). The complex interplay between the social context and the basis for proxies' judgments were also discussed (Chapter 5). Knowledge, beliefs and expectations in the form of trait knowledge, norms and comparisons over time were important anchors upon which proxies based their decisions. Alterations in proxies' cognitive and affective states may also have a bearing on how their judgments change. As discussed in Chapter 6, the experience of living through bereavement can result in various changes in the mental processes that characterise grief. These changes may induce changes in the way events are perceived and judged. Further to this, adaptive psychological mechanisms used to cope with loss could explain changes over time. For example, in Lazarus and Folkman's (1984) theory of stress, the extent to which an event is perceived as stressful is dependent on how it is appraised by the individual. In an effort to cope with the situation, Folkman and Lazarus (1990) suggest that individuals engage in strategies to reduce the threat, such as positive reappraisal and avoidance. Since bereavement over the loss of a significant other is a highly stressful event, there is ample evidence to support Lazarus and Folkman's (1984) theory as it applies to coping with loss (Stein *et al.*, 1997; Folkman, 1997; Moskowitz *et al.* 1996).

## **7.2 Aims**

By investigating the differences and similarities in the accounts given by proxies through their ratings of patients' symptoms using VOICES and through qualitative analysis of the

Proxy Semi-structured Cognitive Interview (PSCI), it is possible to examine whether and how perceptions change over time. This may have a bearing on the use of retrospective measures, such as VOICES, during the bereavement period. If there are significant changes in the reports between different time periods, then this raises questions about the validity of the retrospective approach in palliative care research.

**Research question four**

How do proxies’ perceptions of patients’ pain, anxiety and depression change during the bereavement period?

**7.3 Methods**

As a means to investigate whether and how proxies’ responses change over time a repeated measures design was used. The design and procedures are presented in Section 3.6. Participants who were interviewed at 3-5 months (mean 3.9 months) following the patient’s death (Time 1) were invited to take part in a follow-up interview (Time 2). After completing the Caregiver Strain Index (CSI) (Robinson, 1983), State Anxiety Inventory (SAI) (Spielberger *et al.*, 1983) and the Beck Depression Inventory (BDI) (Beck *et al.*, 1996), a shortened version of the PSCI (contained in Appendix K) was undertaken. Qualitative analysis of the PSCI will be detailed along with the findings in Section 7.4.6.

**7.4 Results**

**7.4.1 Descriptive statistics**

**Demographic and patient/proxy characteristics**

From the sample of thirty participants who were interviewed at time 1, 13 agreed to be interviewed again at time 2. Descriptive statistics of patient and proxy characteristics were produced using SPSS/Win 10.1 and are summarised in Table 7.1 and 7.2. A series of chi-square analyses were conducted to examine differences between those who took part and those who declined. The analyses focused on patient and proxy demographic/characteristics. In some instances, categories were collapsed to enable analysis. Where the frequencies were less than

**Table 7.1: Patient demographics and characteristics.**

| Category | Criteria           | N | (%)  |
|----------|--------------------|---|------|
| Age      | Range 25-86 years  |   |      |
|          | Mean 67 (S.D.13.2) |   |      |
| Gender   | Males              | 6 | (46) |
|          | Females            | 7 | (54) |

**Table 7.1: Patient demographics and characteristics.**

| Category                                      | Criteria               | <i>N</i> | (%)    |
|---|------------------------|----------|--------|
| <i>Ethnic group</i>                           | European               | 13       | (100)  |
| <i>Socio-economic status</i>                  | Class 1                | 3        | (23.1) |
|   | Class 2                | 2        | (15.4) |
|   | Class 3                | 3        | (23.1) |
|   | Class 4                | 2        | (15.4) |
|   | Class 5                | 3        | (23.1) |
| <i>Cause of death</i>                         | Lung cancer            | 4        | (30.8) |
|   | Colorectal cancer      | 1        | (7.7)  |
|   | Upper GI cancer        | 1        | (7.7)  |
|   | Breast cancer          | 2        | (15.4) |
|   | Prostrate cancer       | 3        | (23.1) |
|   | Liver cancer           | 1        | (7.7)  |
|   | Coronary heart failure | 1        | (7.7)  |
| <i>Place of care in the last week of life</i> | Home/hospice           | 4        | (30.8) |
|   | Home                   | 3        | (23.1) |
|   | Hospice                | 3        | (23.1) |
|   | Hospital               | 1        | (7.7)  |
|   | Hospital/home          | 1        | (7.7)  |
|   | Nursing home           | 1        | (7.7)  |
| <i>Place of death</i>                         | Hospice                | 8        | (61.5) |
|   | Home                   | 3        | (23.1) |
|   | Hospital               | 1        | (7.7)  |

**Table 7.2: Proxy demographics and characteristics.**

| Category                     | Criteria           | <i>N</i> | (%)    |
|------------------------------|--------------------|----------|--------|
| <i>Age</i>                   | Range 27-76 years  |          |        |
|                              | Mean 55 (S.D.13.2) |          |        |
| <i>Gender</i>                | Males              | 7        | (54)   |
|                              | Females            | 6        | (46)   |
| <i>Ethnic group</i>          | European           | 13       | (100)  |
| <i>Socio-economic status</i> | Class 1            | 3        | (23.1) |
|                              | Class 2            | 2        | (15.4) |
|                              | Class 3            | 3        | (23.1) |
|                              | Class 4            | 2        | (15.4) |

**Table 7.2: Proxy demographics and characteristics.**

| Category  | Criteria                  | <i>N</i> | (%)    |
|---|---------------------------|----------|--------|
|   | Class 5                   | 3        | (23.1) |
| <i>Relationship to patient</i>  | Partner                   | 8        | (61.5) |
|   | Son/daughter              | 5        | (38.5) |
| <i>Amount of contact between the patient and proxy in the last week of life</i> | 24 hours per day          | 10       | (76.9) |
|   | 12-24 hours per day       | 1        | (7.7)  |
|   | 3-6 hours per day         | 2        | (15.4) |
| <i>Residence</i>  | Lived with patient        | 10       | (76.9) |
|   | Did not live with patient | 3        | (23.1) |

5, which violates the assumptions of chi-square, Fisher’s exact test was used. Since none of the analyses were significant at the level of  $\alpha = .05$ , this indicated that the follow-up sample was not significantly different from those interviewed at time 1 and who did not participate at time 2. (Appendix P contains the full analysis). The test statistics were patient age ( $z = -.139$ ,  $p = .17$ ); patient sex ( $\chi^2 = .14$ ,  $p = .71$ ), patient socioeconomic status ( $\chi^2 = .07$ ,  $p = .79$ ), proxy age ( $z = -1.17$ ,  $p = .24$ ), proxy sex ( $p = .13$ ), proxy socioeconomic status ( $\chi^2 = 22$ ,  $p = .64$ ), place of death ( $p = .69$ ) and proxy residing/not residing with the patient ( $p = 1.00$ ). Amount of contact between the patient and proxy, and ethnicity could not be analysed, due to a lack of variability that resulted in too few numbers in some categories to analyse the data. Place of care consisted of various combinations of care provision again there too few numbers in the categories. The Mann-Whitney test was used to analyse the continuous variables (CSI, SAI, BDI) because the variables were not normally distributed and sample sizes were small ( $N = 13$ ). The Kolmogorov-Smirnov (K-S) two-sample test ( $\alpha = .05$ ) showed that the data met the assumption of the Mann-Whitney test for similarity of distributions; CSI (K-S  $z = .88$ ,  $p = .45$ ), SAI (K-S  $z = .14$ ,  $p = .15$ ), and BDI (K-S  $z = .77$ ,  $p = .66$ ), as Table 7.3 shows, there were no significant differences between the groups on these measures.

**7.4.2 Proxy self-report measures**

Proxies’ ratings on the CSI, SAI and BDI were analysed and compared with their ratings at time 1. Out of a possible maximum of thirteen, over half the sample (8/13) reported moderate to high levels (score  $\geq 7$ ) of perceived caregiver strain with a range of 3-10. The mean score on the SAI (40.5, S.D.13.0) was above average for adult populations of males (mean 35.72, S.D.10.4) and females (mean 35.20, S.D.10.61). Similar to the larger sample, 30% had scores above the normative range for males (46.21) and females (45.81). Scores ranged from 24-72. Reported levels of depression from the BDI suggested that the majority (10/13) of the sample had normal or mild levels of depression (mean 11.5, S.D.11.2). The variance and range (3-43)



**Table 7.3: Analysis of differences between proxies who took part in the follow up (time 2) and those who did not in self-reported levels of carergiver strain, anxiety and depression.**

|      | Non follow up sample |           | Follow up sample |           | Mann-Whitney |          |
|------|----------------------|-----------|------------------|-----------|--------------|----------|
|      | <i>N</i>             | Mean rank | <i>N</i>         | Mean rank | <i>z</i>     | <i>p</i> |
| CSI  | 17                   | 17.0      | 13               | 13.54     | -1.08        | .28      |
| STAI | 17                   | 13.91     | 13               | 17.58     | -1.132       | .26      |
| BDI  | 17                   | 15.03     | 13               | 16.12     | -.336        | .74      |

Two-tailed,  $\alpha = .05$

showed great variability in scores. Three proxies’ scores indicated elevated levels of depression. One proxy had a score (17) indicating mild to moderate depression (16-19). Another proxy’s score (23) signified severe to moderate depression and one had a score (43) denoting severe depression.

Comparisons between proxies’ scores at time 1 and time 2 were made for the CSI, SAI and BDI. Since the assumptions of parametric analysis were violated, the Wilcoxin signed ranks test was used. Using a two-tailed value with  $\alpha = .05$ , no significant differences were observed for the CSI ( $z = -.14, p = .89$ ) or the BDI ( $z = -.197, p = .84$ ). Differences were found for the SAI ( $z = -2.01, p = .045$ ) with greater levels of anxiety reported at time 1 (Md= 47) compared to time 2 (Md= 40).

### 7.4.3 Analysis procedure

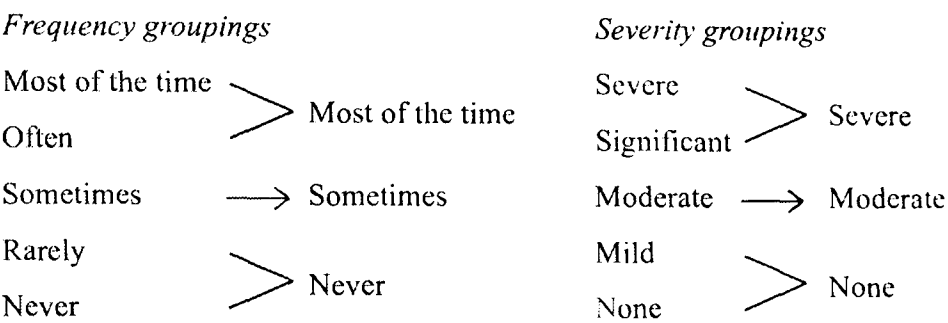
Data from the interviews were analysed in two ways. Proxies’ responses to the VOICES questions regarding patients’ pain, anxiety and depression at time 1 and 2 were analysed quantitatively using SPSS/Win 10.1 and Microsoft EXCEL. The second approach involved using qualitative analysis techniques to examine the proxies’ verbalisations of their thoughts regarding particular questions used in the PSCI. Responses were analysed at both times.

### 7.4.4 Quantitative differences between time 1 and time 2

Cohen’s kappa was used to examine levels of exact agreement in the VOICES responses given at time 1 and time 2. This was chosen because it is designed to control for chance agreement between raters. It was observed that some of the differences were within one rating of a previous rating. Since kappa treats all combinations of disagreements on an equal basis, for instance, a change from *severe* to *mild* is treated the same as a change from *severe* to *significant*, weighting the kappa was considered. This involves assigning a proportion to each and every type of disagreement, so that more conflicting cases affect the measure more than the

less conflicting ones do (Cohen, 1968). Unfortunately, for there to be a valid interpretation, a minimum sample of  $N > 2k^2$  is required (Cicchetti & Fleiss, 1975, cited in Cicchetti, 1976), where  $k$  is the number of points on the continuous-ordinal scale, there are six points on the VOICES scale (including *don't know* responses), meaning a valid interpretation would only be possible if there were greater than 72 subjects. In light of this, a more basic alternative is to group categories based on clinical and empirical findings. From Chapter 4, it is evident that little differentiation was made between some descriptors. For example, *rarely* was often considered so infrequent as to be absent. Likewise, *severe* and *significant* were viewed similarly, as were *most of the time* and *often*. These observations are what provide the basis for the groupings shown in Figure 7.1. As it turns out, grouping is in fact a form of weighting. However, unlike the conventional linear agreement weights for continuous-ordinal scales (Cicchetti, 1976), the weights are nonlinear function of the scale point separations.

**Figure 7.1** Grouping of VOICES frequency and severity descriptors.



Cohen’s kappa values range from 0 (no agreement) to 1 (almost perfect agreement). Assessing the value of kappa is problematic, as varying degrees of acceptability have been reported. Higginson and McCarthy (1993), who were using previous levels as a guide, decided that levels of  $\kappa = .3$  were acceptable. Others have followed Landis and Koch’s (1977) guidelines (Hinton, 1996; Kristjanson, *et al.*, 1998; Lobchuk & Degner, 2002). In keeping with this, benchmarks for the present analysis are kappa values (strength of agreement);  $<.00$  (poor),  $.00-.20$  (slight),  $.21-.40$  (fair),  $.41-.60$  (moderate),  $.61-.80$  (substantial), and  $.81-1.00$  (almost perfect) (Landis & Koch, 1977). Other factors that need to be considered when calculating kappa are bias and prevalence (Byrt *et al.* 1993). Bias and prevalence relate to the balance of the matrix used to calculate kappa. Bias appears when certain disagreements occur more often than others, for example if a large number change their rating in a particular way (from severe to significant and vice versa). While prevalence occurs when participants tend to favour particular responses over others, such as when a high proportion choose *none*. These factors can influence the value of kappa and are considered in relation to the analysis presented here.

7.4.5 Proxies’ responses to VOICES

Pain

Overall exact and grouped agreement for pain was not good. Two proxies stated that the patient did not have pain when they had previously stated that the patient did have pain. For pain frequency in the week, agreement was fair. Almost half (6/13) gave exactly the same response, whereas four of the proxies changed their ratings to less frequent levels. One proxy’s rating became more frequent. In contrast to the last month, exact agreement was moderate for pain frequency ( $\kappa=.45$ ); grouping the ratings increased this to a substantial level ( $\kappa=.64$ ). Here, notable inconsistencies were observed. In particular, one proxy changed their response from *sometimes* to *never*, while another changed from *often* to *rarely*.

Few (5/13) proxies gave exactly the same rating for pain severity in the week and in the month (6/13). Thus, exact agreement at both time periods was poor to fair (Table 7.4). Combining the severity groupings increased the level of agreement in the week ( $\kappa=.28$ ) and month ( $\kappa=.39$ ) to fair agreement. This increase was due to a number of proxies changing their responses from *severe* to *significant* or vice versa. However, there were large variations in ratings as Figures 7.2 and 7.3 show, and as a consequence, agreement was poor.

Table 7.4: Kappa measures of agreement for proxies’ ratings of patients’ pain at time 1 and time 2.

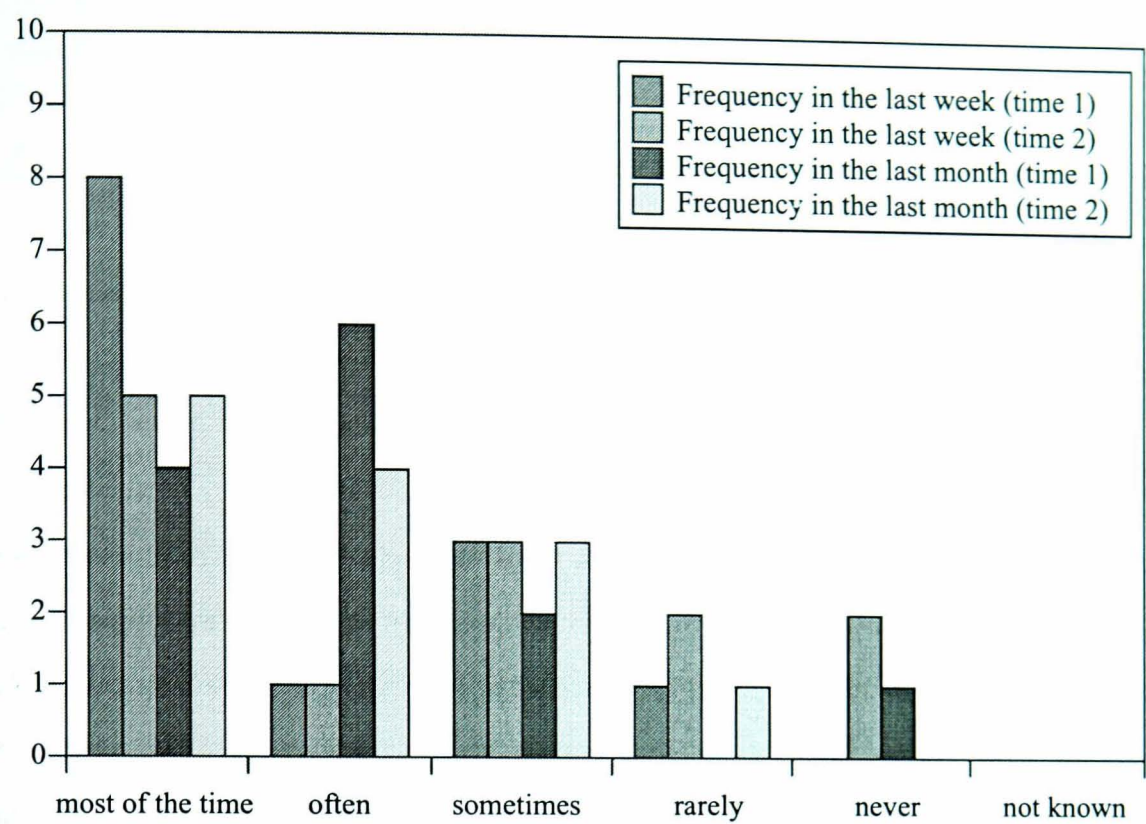
| Reference period | Pain      | Kappa (exact agreement) | Kappa (grouped agreement) |
|------------------|-----------|-------------------------|---------------------------|
| Week             | Frequency | .22                     | .24                       |
|                  | Severity  | .11                     | .28                       |
| Month            | Frequency | .45                     | .64                       |
|                  | Severity  | .25                     | .39                       |

The high proportion of participants answering *most of the time* in the last week of life (Figure 7.2) resulted in prevalence effects in the kappa matrix. The net effect of this was higher marginal totals in the matrix, which translates into a lower kappa value. A balanced spread in the matrix increased the value to  $\kappa=.35$  in this instance. Notwithstanding, agreement remained fair. Similarly, the numbers answering *significant* to pain in the last month (Figure 7.3) showed slight prevalence effects. Here the imbalance in the matrix lowered kappa from  $\kappa=.36$  given an even spread of responses, to  $\kappa=.25$ , still agreement was fair.

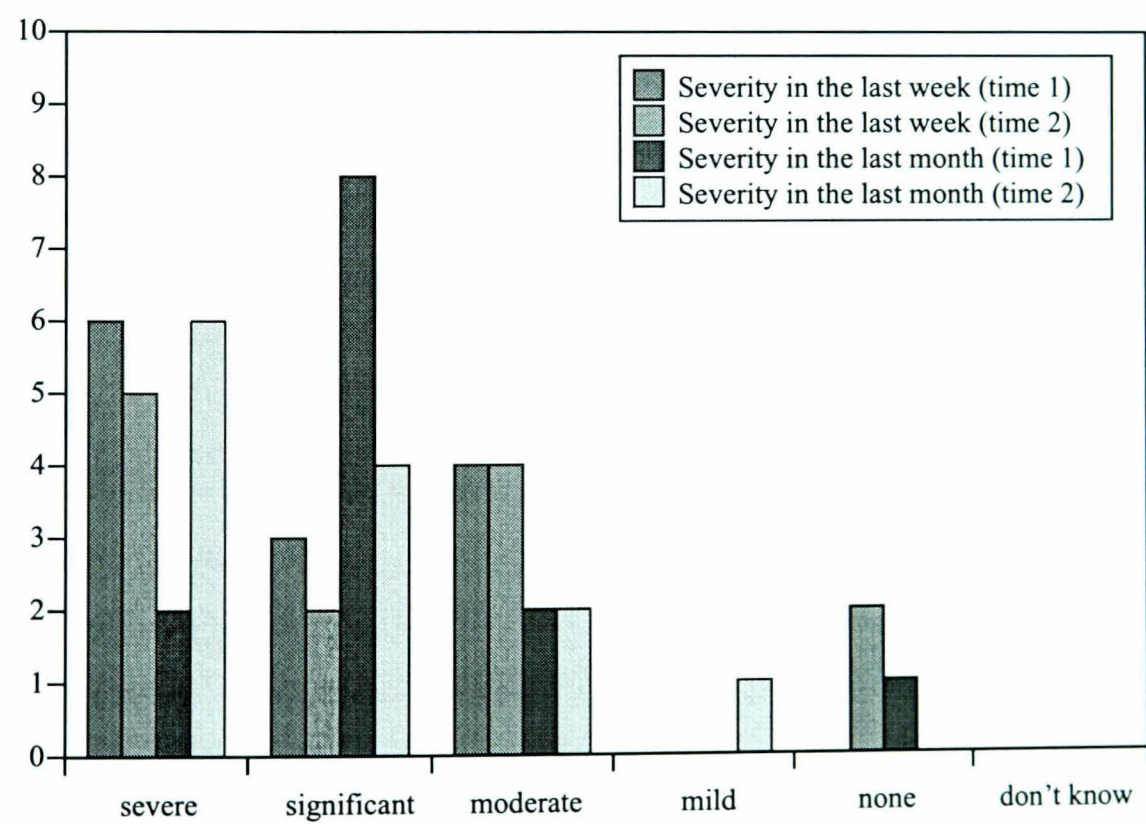
Anxiety

The strength of agreement for anxiety ranged from moderate ( $\kappa=.33$ ) to substantial ( $\kappa=.80$ ). Table 7.5 shows that there was a considerable degree of consistency in ratings particularly for

**Figure 7.2** Proxies' ratings of patients' pain frequency in the last week and month of life time 1 and time 2.



**Figure 7.3** Proxies' ratings of patients' pain severity in the last week and month of life at time 1 and time 2.



anxiety in the month as eleven (11/13) proxies gave exactly the same ratings. Compared to frequency ratings in the month, agreement was lower in the week, where three participants



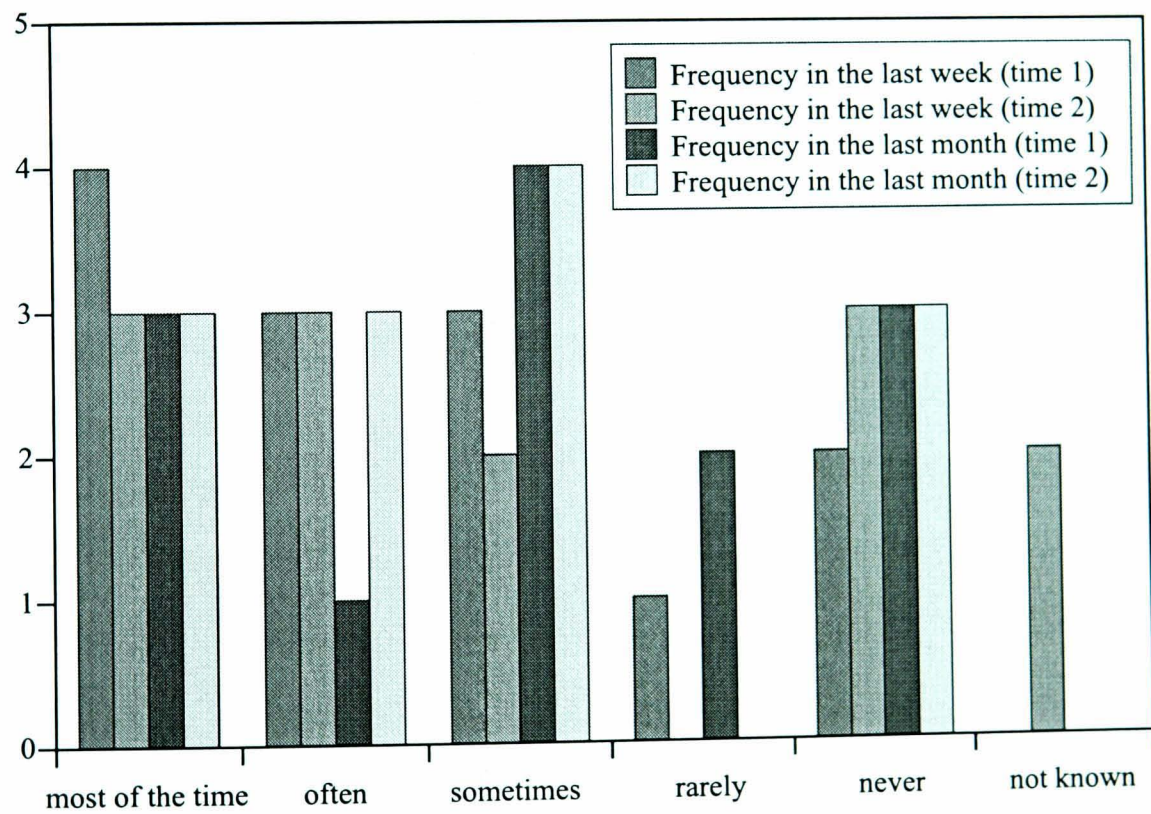
said the patient had anxiety at time 1, and then answered *don't know* or *never* at time 2 (Figures 7.4 & 7.5).

**Table 7.5: Kappa measures of agreement for proxies' ratings of patients' anxiety at time 1 and time 2.**

| Reference period | Anxiety   | Kappa (exact agreement) | Kappa (grouped agreement) |
|------------------|-----------|-------------------------|---------------------------|
| Week             | Frequency | .43                     | .54                       |
|                  | Severity  | .51                     | .64                       |
| Month            | Frequency | .80                     | .77                       |
|                  | Severity  | .57                     | .53                       |

Agreement for anxiety in the week and month were similar, showing moderate to substantial levels. There was little variation in these scores over time. Two exceptions in the week were *severe* and *moderate* ratings at time 1 being rated as *don't know* at time 2.

**Figure 7.4** Proxies' ratings of patients' anxiety frequency in the last week and month of 1 at time 1 and time 2.



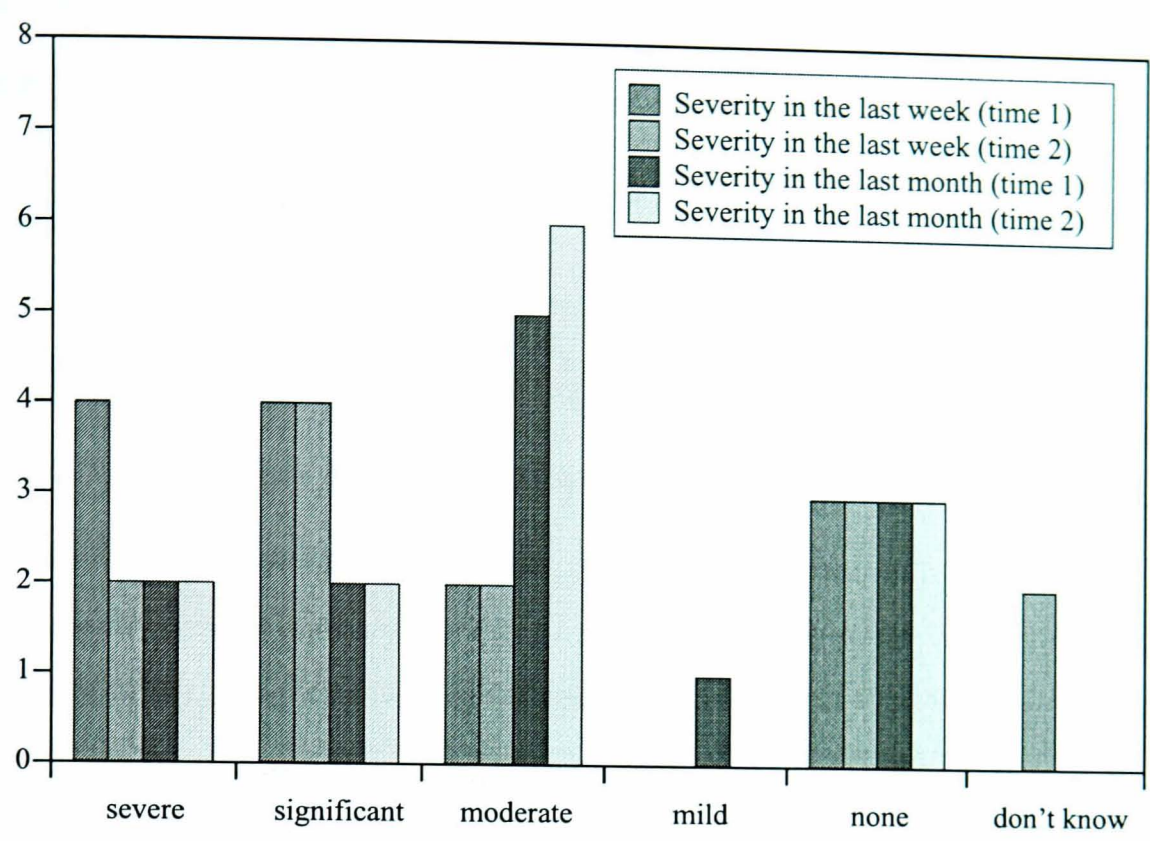
Figures 7.4 and 7.5 show that the spread of responses was more even for anxiety ratings, and consequently, no prevalence was detected.

### Depression

In contrast to anxiety and pain ratings, agreement for depression at time 1 and time 2 were better in the week than in the last month. In the week, moderate agreement ( $\kappa = .48$ ) was observed for depression frequency, while in the month agreement was poor ( $\kappa = .17$ ). Group-



**Figure 7.5** Proxies' ratings of patients' anxiety severity in the last week and month of life time 1 and time 2.



ing the ratings had little effect as major changes were observed (Table 7.6). Four ratings in the week and five in the month changed between *yes* and *no* or *don't know*. In the month, there were both increases (2/13) and decreases (1/13) in ratings of depression frequency (Figure 7.6 & 7.7). As noted, the major changes contributed to the poor levels of agreement for depression severity in the month. Here, three ratings became less severe and one became more severe at time 2. In the week, there was less variation with the two ratings of less severe at time 2.

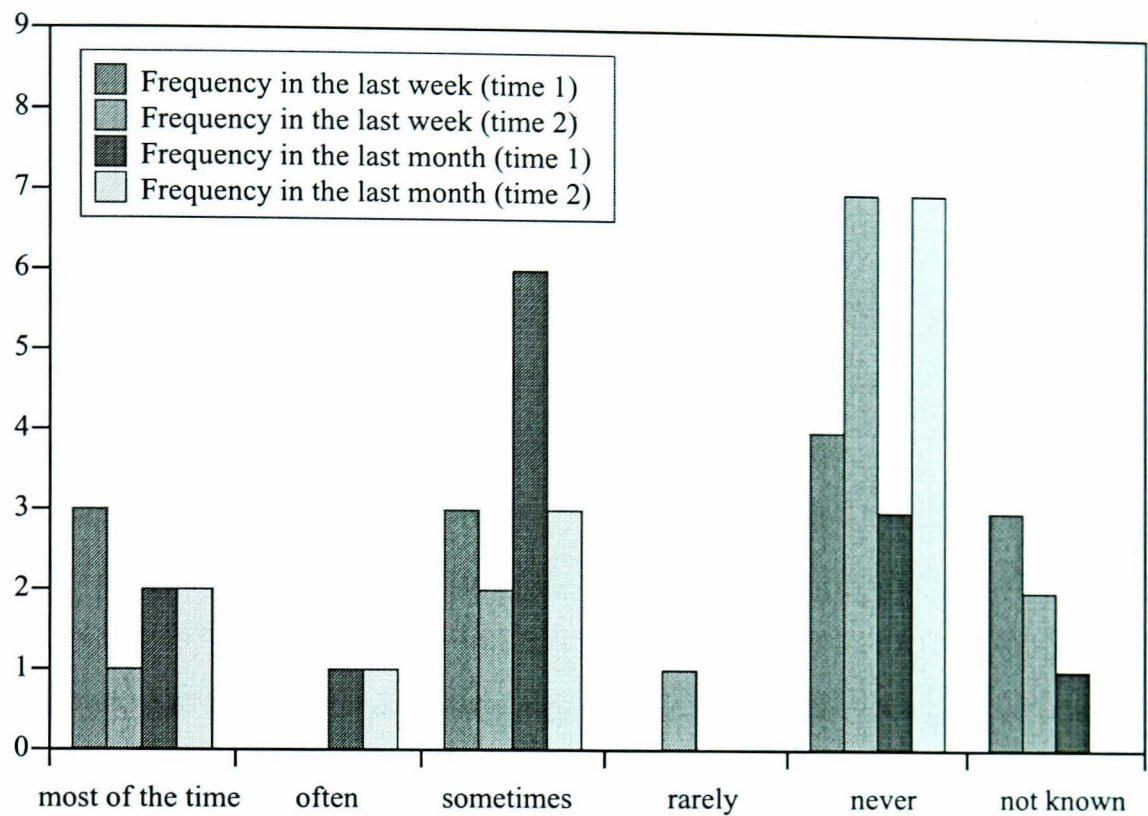
**Table 7.6: Kappa measures of agreement for proxies' ratings of patients' depression at time 1 and time 2.**

| Reference period | Depression | Kappa (exact agreement) | Kappa (grouped agreement) |
|------------------|------------|-------------------------|---------------------------|
| Week             | Frequency  | .48                     | .47                       |
|                  | Severity   | .38                     | .37                       |
| Month            | Frequency  | .17                     | .14                       |
|                  | Severity   | .11                     | .32                       |

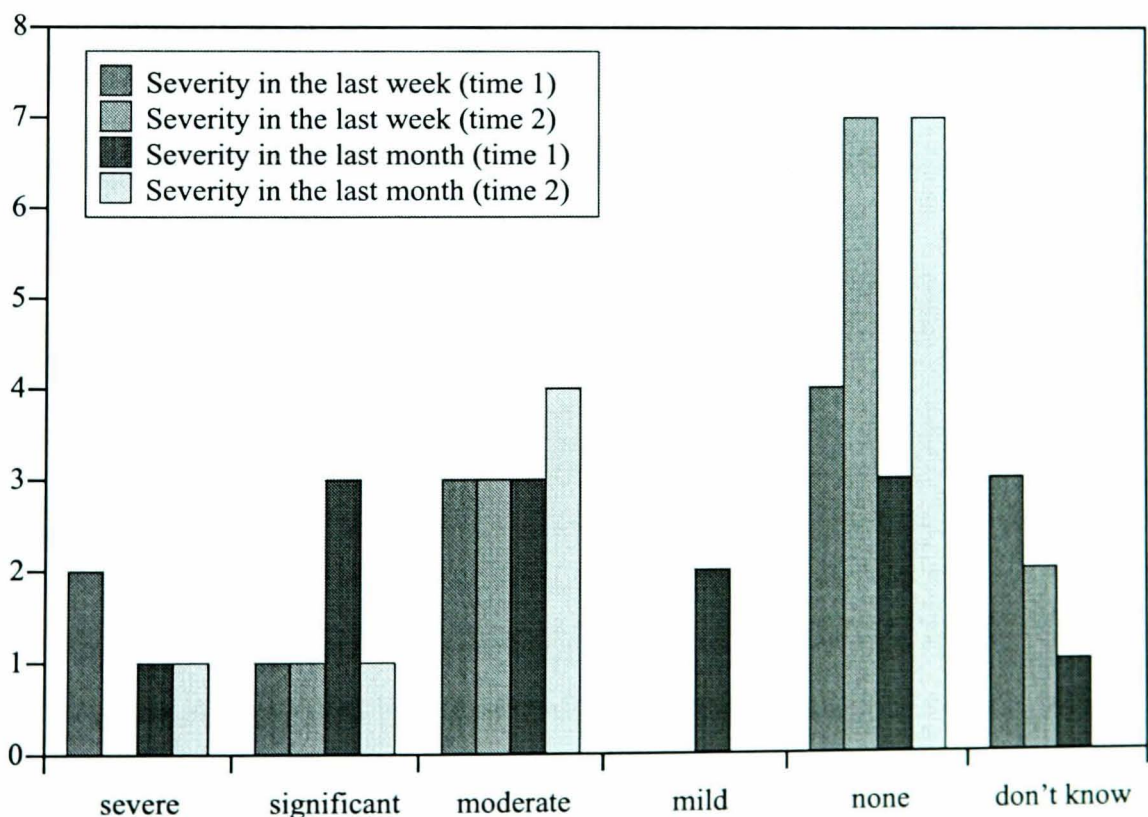
A noteworthy observation regarding proxies' ratings of depression, were the numbers reporting no depression. In the last week of life at time 1, four reported no depression compared to six at time 2. Similarly, in the last month of life, three reported no depression compared to seven at time 2. Due to the numbers reporting no depression, there was some evidence of prevalence. However, the amount was small and lowered the kappa values by no

more than .06 to .09. A hint of bias was observed for depression severity in the week, which had the effect of lowering the kappa value. Altogether, the combined effects of prevalence and bias lowered the kappa values by .06 to .07.

**Figure 7.6** Proxies' ratings of patients' depression frequency in the last week and month life at time 1 and time 2.



**Figure 7.7** Proxies' ratings of patients' depression severity in the last week and month of life at time 1 and time 2.





### 7.4.6 Qualitative analysis

As the questions from the PSCI were structured around the symptoms of pain, anxiety and depression, the analysis was loosely structured into these categories. The analysis differs from the type of content analysis conducted in Chapter 5, where the focus was on the content and frequency of codes as a means to elucidate proxies' question answering processes. The purpose of the analysis conducted more investigative in the sense that the meanings of the utterances were interpreted from the data (Grbich, 1999). All of the texts were entered into winMAX 98Pro, a qualitative computer analysis program to aid the analysis process. The thematic analysis process was as follows:

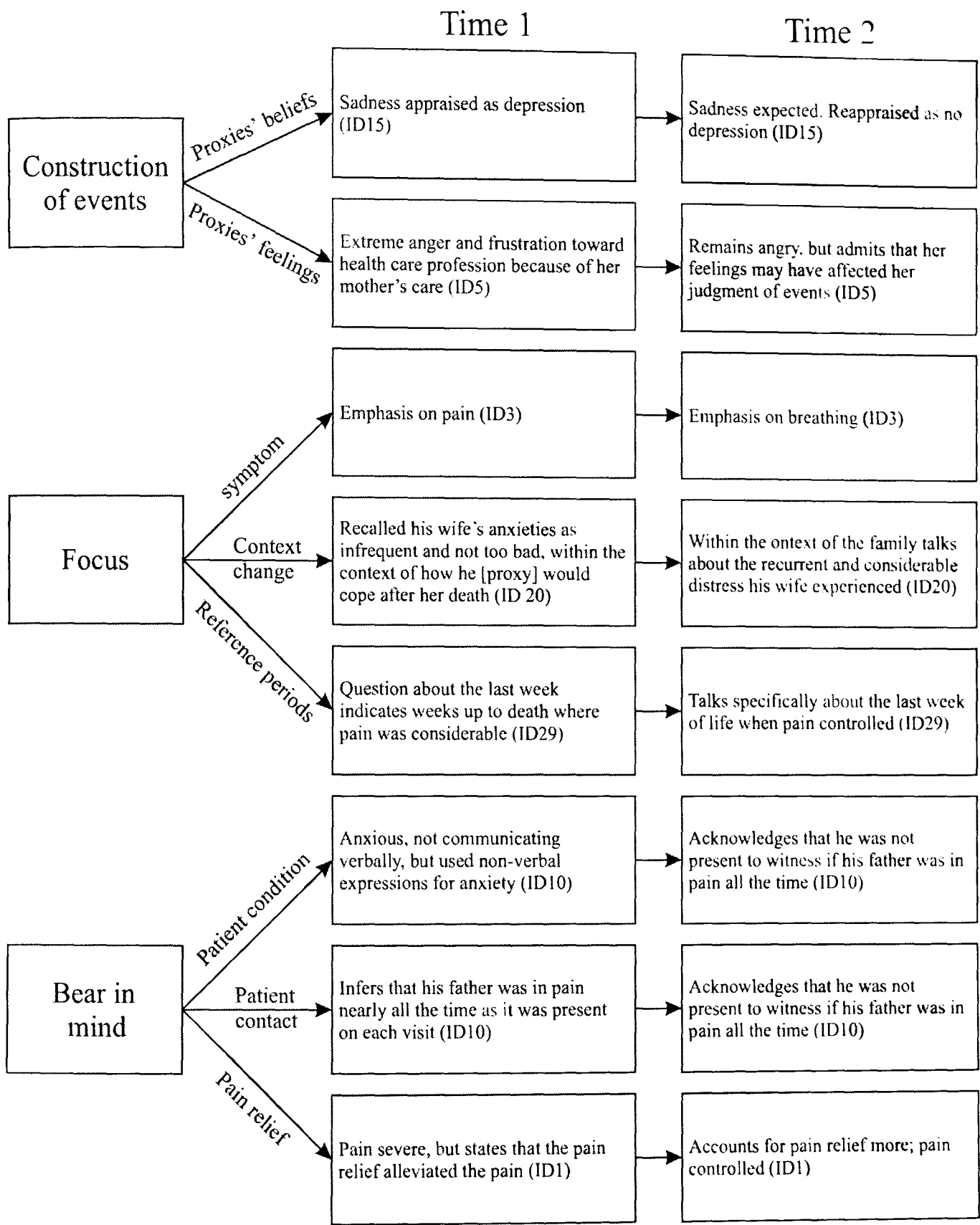
1. Transcripts at T1 and T2 from each of five participants were read and reread several times along with field notes taken at the time of the interviews. Notes regarding the content and initial interpretations were made that characterised differences and similarities in the data. These notes and interpretations served as preliminary themes.
2. The themes identified from the analysis of the first few transcripts was used to guide analysis of further transcripts. New themes were added as they emerged.
3. Associations between similar themes were made, and themes were modified or combined as the analysis progressed. Initially, the symptoms were used as a rough guide to organise the process of analysis, though it became apparent from the themes and patterns that emerged that they went beyond these categories.
4. To ensure the interpretations were valid, themes were continually compared with the transcripts at T1 and T2 and between transcripts from different participants.
5. Some interpretations represented broader categories from which more specific themes were identified. These categories represented themes with related characteristics (e.g. *bear in mind*) or an underlying construct (e.g. *construction of events*). These are discussed in more detail below.

The main categories and themes are presented in Figure 7.8, along with an example of the themes from the interviews. The analysis was conducted separately from the VOICES responses, to reduce bias in coding the interviews. However, within the responses some proxies made reference to their ratings, which could not be avoided. Since a major aim of the study was to examine whether and how proxies' responses changed over time and how this related to their ratings of patients' pain, anxiety and depression, the qualitative analysis is discussed in relation to the VOICES ratings is unavoidable.

***Construction of events*** was a major category that emerged from the proxies' responses. It further subdivided itself into the themes of *proxies' beliefs* and *proxies' feelings*. Within the context of the retrospective interviews, this category referred to changes in the way similar stories were viewed by proxies at different times. The similarity in the responses given by



Figure 7.8 Categories and themes from qualitative analysis.



some proxies was astounding, given the timing of the interviews relative to the patient's death. Many recalled very detailed information about events. However, in several instances, perception of events did appear to change over time. When this occurred the theme proxies' beliefs emerged. In this case, changes in beliefs about the patient's ability to cope with events, norms and expectations were often identified as a possible source for a change in proxies' perspective of events. An illustrative example comes from a proxy talking about depression in

the month before her husband died. The responses were similar at both times, with the proxy describing her husband's sadness and change in mood. At time 1 she recalls this as a "*moder- atish and noticeable*" depression. At time 2 on the other hand, she makes references to her husband's resilience under the circumstances, "*I know he was frightened but I don't think he was depressed...He was extremely brave*" she adds, "*I was surprised how well he handled it*" (ID 2; 24b). In this example, the VOICES rating changed from depression to no depression at time 2. Overall, comparison of this theme with the VOICES ratings showed that the ratings changed almost equally to become either higher or lower ratings of symptom intensity and severity. Talking about his wife as "*a coper*", one proxy talked about the increasing need for medication and how well she coped with the significant pain she experienced. At the follow up interview, he again refers to his wife's ability to cope but states that she was not coping so it must have been severe, as she had to resort to medication (ID 20). There were exceptions to this ( $N=3$ ), where the proxies' perspective and story were completely different from their earlier responses. A good example of this was from a proxy talking about his father's pain. At time 1, he mentions the distress of seeing his father's behaviour change because of the pain, "*towards the end he started swearing and he never swore*" he adds "*I saw a major change in Dad*". In contrast, at time 2 the pain is observed through his father's facial expressions and movements and no reference is made to his earlier behaviour (ID 24).

The second theme in the category perspective was proxies' feelings. Although proxies' feelings were an integral part of many of the responses, this theme included those proxies who acknowledged that it might have influenced their responses. A proxy talking at time 2 about his earlier interview provides an illustrative example of this,

*"I suppose her pain was more controlled than I first said to you. I just saw it as total pain. It's all out of perspective then."*

(ID 1; 10b)

In this example, the proxy's rating of the patient's pain became less severe in the last week of life. Another proxy talked about how detached she had felt from events "*Before [time 1] it was like I was talking about someone else*" (ID 29; 26b). She added that, over time, she had thought about things more and had realised how much her husband had managed to cope with over the last few weeks of his life. Despite greater emphasis on these aspects, in her discussions of anxiety and depression in the last month of his life, her ratings of these symptoms to the VOICES questions did not change. Several other proxies expressed their feelings such as anger, guilt and anxiety but made no reference to the effect this had on their responses. These instances were discussed in Chapter 6.

As a category, *focus* represents a change in central elements within the two time periods. This was evident as a *symptom* focus change from one symptom to another or a greater or

decreased emphasis on the symptom of interest. By way of illustration, when asked about pain at time 1, one proxy talked about the discomfort and pain experienced by her husband, whereas at time 2, when answering the same question the focus was on his breathing (ID 3). Similarly, others talked about problems such as water retention, constipation and itchiness, which were perceived to be particularly difficult to control. Comparing these observations with the VOICES ratings, revealed that in these instances, ratings of the symptoms tended to be less severe or frequent ( $N=4$ ). For example, at time 1, although the proxy had mentioned that itchiness was a problem, the emphasis was on the patient's pain. While at time 2 the proxy states,

*"I would say that the itching was very distressing rather than the pain like, that wasn't a problem."*

(ID 6; 7a)

Since the pain was not considered a problem, the proxy changed his VOICES ratings from *moderate pain most of the time* at the first interview to no pain at the follow up interview.

*Reference periods* included a focus change across the time periods from the last week or month or a period that overlaps or extends beyond these times. For example, talking about her husband's pain in the last week at time 1, one proxy (ID 29; 1b) spoke at length about the severity and frequency of the pain over the last few weeks of his life. At the follow up interview (time 2) the proxy specifically talked about the last week of his life, when the pain was more controlled and only one incident of pain was recalled. In this case, VOICES ratings of pain in the last week of life changed greatly from *severe pain most of the time* (time 1) to *moderate pain rarely* (time 2). This theme coincided with major changes in responses and VOICES ratings across time, which is understandable given that proxies are talking about different time periods. There was no consistency in the direction of changes with some more severe and frequent and other less so. However, it was noteworthy that on whichever occasion the rating was kept within the reference period the rating was lower.

Another theme within this category related to reference periods was *reminiscence*. Here proxies' responses changed from an emphasis on symptoms in the period assessed to periods in the illness trajectory where the symptom was less of a problem to the patient. For instance, at time 1, one proxy (ID 6; 19b; 24b) spoke of his wife's depression in the week and month she died and her sadness at leaving the family. At time 2, when talking about depression, he mentions several anecdotes from when she was diagnosed up until her death where his wife managed to cope well with the illness and maintained her sense of humour. In this case, VOICES ratings changed from *moderate depression most of the time* in the last week and *significant depression sometimes* in the last month, at time 1 to no depression at the follow up interview. Periods often mentioned in this theme were prominent events such as holidays.

anniversaries and family gatherings. This theme was not consistently associated with changes in VOICES ratings. In fact in many instances where this was observed, the ratings remained the same.

***Bear in mind*** as a category incorporated three themes representing changes in proxies' responses as a result of considering factors in one interview that were not considered in the other. *Patient's condition*, as a theme, were changes as a result of the patient's condition being considered. This incorporated the patient's level of consciousness or other conditions such as confusion, or an overlap in symptoms. In all instances where this theme was identified, the change across the time periods moved towards less certainty in the proxy's response ( $N=14$ ), with only one exception where the opposite was observed. Thus, at time 2, responses tended to change from *yes* to *don't know* or *no* responses ( $N=9$ ), while in two instances VOICES ratings stayed the same. A second theme within this category was patient contact. This occurred in two cases where responses at time 2 took into account that contact was limited to visiting the patients in a hospice or hospital. Talking about his father's pain at time 2 one proxy stated,

*"I went up there [hospital] everyday, he was in pain. You could tell from the way he was moving that it was hurting you know. Again it's hard to tell really because outside of the times I was visiting I couldn't really say."*

(ID 10; 2b)

In this case the proxy rated the VOICES pain as *sometimes*, whereas at time 1, the amount of patient contact was not considered and the rating was *most of the time*. In both cases, symptom frequency ratings became less often during the time the patient was in the hospice. Pain severity ratings also became less severe in one of these cases. However, there was one exception with anxiety rated more severely, and this was because the patient was told they were going to die during their stay in the hospice (ID 15). Another theme that emerged, and that tended to lower ratings, was pain relief. At time 1, five proxies' ratings did not take into account the effects of pain relief when asked about pain. This was verified with them during the interviews. This is significant because a number of proxies (at time 1) talked about the patient being in intense pain despite admitting that the pain relief was effective.

### ***VOICES ratings***

Comparing proxies' responses with their VOICES ratings was useful for identifying instances where proxies recalled similar stories and perceived the events in a similar way while changing their VOICES ratings. In every such case, the ratings did not fit into the categories identified. In almost all instances, the change was by one rating. The most commonly occurring was a change between *severe* and *significant* ( $N=6$ ). Another was between *most of the time* and *often*. A possible explanation for this was the similarity in the descriptors. This was borne out

in several of the responses where symptoms were often referred to as *severe* on both occasions but the rating changed between *significant* and *severe*. In two responses, the proxies explicitly stated that the symptom was present almost all the time at both interviews, but rated VOICES as *most of the time* on one occasion and *often* at another. In far fewer cases, was the rating changed from *rarely* to *none* ( $N=1$ ). The findings add support to the grouping of the VOICES descriptors as described in the quantitative analysis above and the findings in Chapter 4.

### ***Reliability and validity***

The analysis conducted in this chapter focused on changes in proxies' responses over time. Although references are made to the frequency of themes and categories, the emphasis was on describing and interpreting changes. Both field notes and transcripts of the interviews were vital in capturing a valid interpretation of what the proxies were saying. As the sole interviewer, having contact and knowledge of proxies' VOICES ratings may have inadvertently created bias in analysing the data. To reduce this during the analysis, the transcripts were kept separate from the VOICES ratings. As mentioned earlier, knowledge of the ratings became evident in some instances as the proxy in their responses mentioned them. Unfortunately, this could not always be avoided. In order to reduce possible biases and to ensure that the interpretations were valid, comprehensive and reliably applied across all of the transcripts, the categories and themes were discussed with another researcher. Any differences were debated until agreement was reached. In some cases, it was also possible to verify some themes with the proxy at the time of the interview (respondent validation). In particular, pain relief, contact, patient condition and proxies' feelings. To avoid undermining their responses, no other themes were verified with the proxy. Consistency of themes was also established by going back over some of the transcripts several months after the analysis to examine whether the interpretations and themes identified were the same. The analysis was comprehensive as it was based on all the transcripts and not just on excerpts.

As a result of the analysis of the data at time 1 and time 2, it is apparent that there are instances where the changes in VOICES ratings do not coincide with any of the identifiable themes and categories or were not consistently associated with changes. These anomalies may be attributed to the reliability of the VOICES instrument or other factors that could not identified through the methods employed. It should be noted that proxies might not always be aware of the factors that are likely to influence their responses. For example, the theme proxies' feelings included responses that showed an awareness of that feelings were having an effect on their own responses. These feelings may influence judgments without conscious awareness. The present interpretation relied on identifying and attaching meanings to consist-

encies and inconsistencies in proxies' accounts. In a minority of instances, the responses and VOICES ratings were completely different. There was no way of identifying why this occurred.

## 7.5 Discussion

Through their responses, proxies revealed the complexity of cognitive processes underlying their judgments at each interview, and across time through comparison of their interviews. Analysis of the VOICES ratings was a means to specifically identify consistency and disparity in responses, whereas, qualitative analysis of the PSCI increased awareness of factors likely to influence change. From the VOICES ratings, varying levels of agreement were found depending on the symptom, the reference period and the aspect assessed (frequency, saliency). Overall, anxiety ratings showed a considerable level of agreement in the week and month. This was evident during the interviews, where in many instances; patients' anxieties were associated with particular events, such the last rites, symptom control, patient diagnosis and hospice admission. These events were especially vivid and recalled with incredible detail and consistency over time, supporting the memory literature regarding saliency and memorability (Bower *et al.*, 1979; Christianson, 1992). When differences emerged, ratings tended to be lower at time 2. One explanation is that proxies are more likely to state that they were uncertain (*don't know* responses) due to the patient's condition or their contact with the patient. This was identified in the *bear in mind* category. This was also exhibited in both pain and depression. Other inconsistencies in all symptom ratings coincided with the category focus, with differences in the reference periods assessed. This highlights the need to continually reiterate to the proxy the period being assessed to confine their responses accordingly.

Overall, pain and depression ratings were not as consistent as anxiety. Pain ratings in the week were disparate, which resulted in poor to fair agreement. Accounting for slight differences in ratings by grouping variables did indicate better agreement for pain in the last month of the patient's life. Apart from the patient's condition, another explanation for the uncertainty and changes in responses in the last week might have been the proxy's level of involvement in the patient's care. Since many of the patients were cared for in the hospice during the last week, where previously they had been at home with the proxies involved in their pain management. In contrast to pain ratings, depression ratings were more consistent in the week than the month. The numbers reporting *no* or *don't know* in the week remained consistent. Most differences occurred because of uncertainty at time 2, when proxies had previously stated that the patient had depression. Discrepancies in pain and depression ratings tended to coincide with a variety of themes observed in proxies' responses.

Proxies' beliefs, under the category *construction of events*, corresponded with both increases and decreases in VOICES ratings for pain and depression, although ratings did not always alter. Changes in the way similar events were viewed was a major category. Here, differences in proxies' beliefs regarding patients' ability to cope, norms and expectations were apparent from their responses. Norms and expectations were more likely to be expressed for depression, where sadness was to be expected rather than depression. For pain, the patient's ability to cope was mentioned more often. Combined with the theme *reminiscence*, these results may indicate an attempt by the proxy to deal with the loss by reappraising the situation, or by thinking back to less distressing times (Lazarus & Folkman, 1984; Taylor, 1991). Individual differences in adapting during bereavement may explain why some VOICES ratings diverge over time. Alternatively, it is also possible that proxies may have more opportunity to talk to other bereaved individuals and may undergo events during the intervening period between interviews. These, in turn may have a significant influence on their perception of events. Anecdotal evidence from proxies interviewed suggests that several had been in contact, through hospice support groups and social support networks with other individuals who had recently experienced bereavement.

A change in the focus of events, highlighted in the category *focus*, was more notable for pain, with an emphasis on change to other symptoms perceived by the proxy as more troubling to the patient. The poignancy of these symptoms may have remained more vivid over time. Further to this, the interview itself may have prompted proxies to express these problems within other questions, as the focus was very precise not providing evaluation of symptoms other than pain, anxiety and depression.

The present investigations highlight the importance of timing in gathering information from proxies retrospectively. Some measures within the survey can be put in place to reduce some changes over time. For example, the category *bear in mind*, included themes that could be integrated into the design of surveys, such as accounting for pain relief, amount of contact and reference periods. Changes in proxies' internal states during bereavement are more problematic for the retrospective approach. What is important is the accuracy of these reports in relation to the patient's experiences. In these investigations, it was not possible to assess which account (time 1 versus time 2) was more valid or whether either actually reflected events. This requires further inquiry. Ideally, it would be necessary to compare patients' accounts with those of their proxies at different periods throughout bereavement. This would be useful for determining the most suitable time for interviewing. Overall, these investigations do indicate that, in many cases, ratings become less severe and less frequent with the passage of time. In addition, proxies became less certain about whether the patient had experienced the symptom at the follow up interview. This shares similarities with Cartwright and

Seale's (1973) study where proxies were less certain at 9 months compared to 3 months. However, this was in relation to place of death.

Significant differences in some proxies' responses with the passage of time do raise important methodological questions with the retrospective approach for patients at the end of life. As this exploratory investigation has shown, there is consistency for some symptoms (e.g. anxiety) over certain reference periods. Given the importance of this method in the audit of patient care and the lack of research into the validity of proxies, researchers cannot afford to dispense with it at present. Instead, future research should focus on determining the most appropriate reference period, symptoms and timing. The findings from this study are a useful starting point from which to generate hypotheses that should improve the validity of the approach.



## CHAPTER 8: Discussion

Care at the end of life is important for meeting the physical, social, psychological and spiritual needs of patients and their families. However, evidence suggests that there are significant shortcomings in the provision and planning of these services (Addington-Hall & McCarthy, 1995a). Audit and evaluation are vital to ensure that patients and families receive appropriate care. Given the methodological problems associated with evaluating the care of the dying, investigators have turned to alternative sources for information on patients' experiences, such as bereaved relatives. The systematic review in Chapter 2 revealed that a limited number of studies have addressed the validity of retrospective proxy reports in palliative care. In those studies that have, small sample sizes and conflicting time periods limit the conclusions that can be drawn from them. By examining the literature from both prospective and retrospective proxy validity studies various factors that influence agreement between patients' and proxies' accounts were revealed. Agreement on subjective aspects of patients' functioning was found to be poor. A pattern that is highlighted repeatedly in proxies from other areas of health (Chapter 1).

The motivation for this investigation hinges on the fact that there is poor agreement for these aspects of patients' functioning. Since pain, anxiety and depression are by nature subjective and of critical importance in palliative care, they were the focus of the present investigation. Despite finding several factors that contribute to the level of incongruency observed, none of the studies reviewed in the aforementioned chapters had specifically investigated what proxies were actually basing their judgments on. Since there was little evidence on the validity of proxies after the patient's death, this Ph.D. endeavoured to identify the current shortcomings and understand the possible reasons for them as a means to improve measures. Four research questions embodied the aim and form the basis of the present work.

### 8.1 Research aims and findings

#### 8.1.1 Research question one

Research question one set out to investigate the descriptors used in VOICES for describing the frequency and severity of patient pain, anxiety and depression. The purpose was to examine how proxies' interpreted the terms, identify any ambiguities with them, and assess their validity. By analysing the proxies' responses to the severity descriptors, a number of behaviours, feelings and thoughts were identified. In some instances, such as for high levels of pain, the participants had a consistent view of the term. Many of them described it as sharp, prolonged, distressing and uncontrollable. At lower levels, pain was regarded as more of an ache that was manageable and periodic in nature. Not surprisingly, the category feelings/cognitions

was predominant for anxiety and depression. Proxies' descriptions showed a good awareness of anxiety and depression symptomology, even though none of the participants interviewed had any formal medical training. There was also evidence that some proxies could distinguish between anxiety and depression. A reduction in functioning, interest in activities and mood negativity characterised depression. In contrast, functioning was hardly ever mentioned for anxiety, instead awareness of concerns prevailed. Like proxies, pain severity descriptions, persistence of anxiety and depression was associated with higher levels, whilst sporadic episodes were commonly described at lower levels.

This investigation identified several problems with the VOICES descriptors. For the severity verbal rating scale, the term *significant* was poorly understood by participants. This was evident in their diverse interpretations and their tendency to use other descriptors on the same scale to explain its meaning. To avoid ambiguity associated with this term, it should not be used<sup>3</sup>. Another term that caused problems was *rarely* in relation to frequency of anxiety and depression. When used in conjunction with a short assessment period, the term is almost synonymous with *never* and is therefore superfluous. Over twelve months, this may not be a problem. A possible solution is to adopt proxies understanding of the term "*rarely*" by replacing it with "*occasionally*" or "*occasional episodes*". The lack of differentiation between *most of the time* and *often* also needs attention, as the term *often* is redundant. One suggestion would be to change *most of the time* to *all of the time* and retain *often*. This would effect a separation between the two points on the scale. It should also be noted that several proxies stated that their relatives were in pain *all of the time*, thus lending support for its inclusion. The descriptor *sometimes* is another term whose meaning was vague. In spite of the uncertainty in its precise meaning, it was generally perceived to be the middle point of the frequency scale.

The recommended changes may reduce some of the diversity in the interpretation of the descriptors. The authors of VOICES need to decide on the level of sensitivity they require, as some variations in interpretation of the descriptor may not be problematic. For instance, is it necessary for the scale to distinguish between daily episodes of pain and several episodes of pain in per week?

### 8.1.2 Research question two

Based on the findings arising from research question one, research question two sought to understand what proxies were recalling when they actually used VOICES to answer questions about their relatives. This is important if measures to increase the accuracy of proxies' accounts are to be implemented successfully, incongruency and biases need to be better

---

3. More recent versions of VOICES no longer use this term.

understood. The areas explored were responses difficulties, strategies upon which decisions are made, cues used to form the basis for judgments and reference periods.

The findings indicated that proxies experienced response difficulties specifically when asked to assess patients with terminal illnesses. Difficulties were evident when the patient's level of consciousness decreased, which was more likely in the week before death when the their health deteriorated and/or because of symptom control measures. This made it difficult for the proxy to assess the patient's symptoms as there were few cues, either verbal or non-verbal from which to draw upon. A lack of communication between the patient and proxy also presented some difficulties, particularly in assessing anxiety and depression. Discussing psychological distress may be very difficult for the patient and proxy, given that death is imminent. The findings suggest that proxies are aware that patients often do not want to discuss their emotions, and consequently they compensate for this in their judgments. Anxiety and depression are very difficult to differentiate, because the symptoms are very similar. However, this did not appear to be a problem for most participants, although a small number did talk about anxiety and depression being the same thing.

Despite the fact that there appears to be a preponderance of possible response difficulties, it is important to highlight that the vast majority of proxies did not encounter any. For those who did, measures need to be included in the design that ensure that those proxies willing to give a response, despite their uncertainties, be excluded. Under these circumstances, proxies' responses are likely to be inaccurate and may affect the outcomes of the study. Proxies should be given the option, within the survey, to state that they are unable to judge the patient's symptoms. For example, *"If you are uncertain about whether your relative had pain or not in the week before they died because of their condition, please indicate"*. Although, the amount of data would be reduced, its validity would be improved. Such a trade-off is vital if researchers are to gather credible data.

The capacity of this study to identify the heuristics or rule-of-thumb methods used by proxies was limited because of ethical concerns for the participants. The estimation strategies employed by proxies to respond to frequency questions may influence levels of agreement if patients and proxies use different strategies. One possible solution is to phrase the question in such a way that the strategy is evident. For example, if proxies are asked to assess a symptom by averaging episodes over a specified reference period. This would introduce some consistency into the responses and might improve congruency between the patient and proxy. Proxies' judgments about symptom severity were also found to vary. Some proxies chose the most severe level or the most frequent level, while others where not able to verbalise their strategy at all. It is recommended that the question contain all of the information that is required, including any constraints. For example, *"on average over the last month, how severe was his/*

*her pain” or “over the last week, what was the most severe level of pain he/she experienced?”.*

Various cues were used by proxies to judge the presence, frequency and severity of the patients’ symptoms. Decisions about pain were based on three main indicators, namely, non-verbal and verbal communication, and medication usage, and two minor ones, functioning and control. In Chapter 4, the latter indicator was much more important, and was used by the majority of proxies. This is likely because control is a more perceptual experience that is hard to discern without other clues. Like control, anxiety and depression are, in essence, subjective experiences. However, the findings imply that regardless of this, proxies are able to identify what they thought were anxiety and depression through various more overt indicators. Anxiety was inferred from patients’ verbalisation of concerns, behaviour and functioning in the month. Cues for depression included verbal and non-verbal communication, as well as disruption to functioning. By having knowledge of the patient’s usual demeanour (trait knowledge), proxies were able to compare it with their present mood state and/or changes that occurred over the disease trajectory. Since health professionals do not have access to such detailed knowledge about a patient, they may have to rely on other indicators. This may account for the differences between health professionals’ and significant others’ ratings when compared to the patients’ self-reports. For significant others, this knowledge may be invaluable for making accurate judgments about the patient’s psychological status. Clearly, more research is needed to compare the indicators used by proxies with the patient’s actual symptoms so that the reliability of the cues can be assessed. There is evidence from research using third parties to assess mood states in patients with Alzheimer’s disease, that various indicators can be used to detect anxiety, depression and anger. For example, the Apparent Affect Rating Scale relies on observations of the patient’s facial expression, voice quality, body movements and verbalisations (Lawton 1996). Overall, the findings provide evidence that, in judging patients symptoms, proxies integrate several pieces of information in sophisticated ways.

From the analysis most proxies kept within the reference periods for pain in the week and month and for anxiety in the week. Even where reference periods were extended (anxiety in the month and depression in the week and month) the numbers were relatively small. However, in a large survey this might constitute a considerable number. As discussed in section 5.5.8, where *forward telescoping* occurred it was usually the result of the proxy discussing a salient event, such as the patient’s diagnosis, treatment or a deterioration in the patients condition, which coincided with the patient becoming anxious or depressed. Apart from an emphasis on the reference period within the survey, other measures to help control for this were presented in Section 5.6. For example, using the proxies personal frame of reference to aid accurate recall.

### 8.1.3 Research question three

The retrospective approach has been criticised primarily because information is gathered when the proxy is going through the grieving process. This is thought to account for many of the differences observed in ratings taken before and after the patients death (Higginson *et al.*, 1994; Hinton, 1996). In Section 6.1, it was noted that few studies have actually examined the effect of proxies' psychological distress in relation to congruency between patients and proxies' reports. The aim of Chapter 6 was to explore whether proxies' own feelings formed part of the basis for their judgments, and in so doing, influenced their ratings of patients' pain, anxiety and depression using VOICES.

The recognition that proxies' own feelings were included in their responses was not unexpected given the types of relationships and the amount of contact they had with the patient. Many experienced and shared anxieties about symptom relief and patient care and death. Although it is not possible to ascertain the influence of proxies' feelings on the accuracy of their responses in the present investigations, it is clear that their feelings formed part of their thinking when judging patients' symptoms. This was supported by the identification of affect-based judgments in their responses, where their own concerns and feelings were used as a basis to judge patients' symptoms. An example was knowledge based on similarities between patients and proxies, or an intuitive sense of the patient's experiences in long-standing relationships. In some cases, similarities may have provided an accurate representation of patients' experiences, but this cannot be confirmed. Since the situation is also novel to both, the patient may not respond as expected and this may lead the proxy to erroneous judgments.

In light of the finding that there is a propensity for those with greater caregiver strain and anxiety to express their feelings and use affect-based judgments, and for the significant differences found for some of the VOICES ratings (Section 6.4.2), it is necessary to take these factors into account when using information derived from proxies. Although feelings and affect-based judgments tended to occur in proxies with greater caregiver strain and anxiety, the results were non-significant. More research is required to explain the underlying processes. Further investigation is also required in order to ascertain the affects of anxiety on proxies' ratings, so that methods to account for them can be implemented. If there is no systematic bias, one suggestion put forward by Magaziner (1992) is to increase sample sizes such that the standard error of judgments is minimised. Where there is bias, as in the case of caregiver strain, the direction and magnitude should be determined so that measures, such as weighting, can be used when interpreting the results (Magaziner, 1992). Some prospective studies have started assessing the degree of bias (Sneeuw *et al.*, 1997a; 1997b; 1998; 1999; 2001; Lobchuk & Degner, 2002a). So far findings from these studies suggest that the magnitude of bias is small to moderate.

### 8.1.4 Research question four

Another gap in the knowledge is the reliability of proxies' responses over time, which have obvious implications for the audit of care and the validity of the retrospective approach. For instance, if proxies' responses change significantly, how do they change and when is the best time to interview? Previous research comparing proxies' responses before and after the patient's death has shown that they are liable to change. With the exception of the present work, no studies included a follow-up on the same group of proxies during bereavement. Thus, there is no prior work with which to compare the present findings. Research question four was formulated in response to this omission in the proxy validity literature. It set out to examine whether and how proxies' perceptions change over time.

By comparing VOICES ratings at two different time periods, a number of inconsistencies were revealed. Possible reasons for these differences were extracted from a qualitative analysis of the responses to the same questions at time 1 and time 2. As discussed in Section 7.4.6, there are some similarities and differences. For the week reference period, pain ratings were found to have poor to fair agreement, even accounting for the slight changes in ratings by grouping them. Agreement was better in the month as it ranged from fair to substantial. There are several explanations for these findings. Firstly, the patients condition in the week created greater uncertainty at time 2, with more *don't know* and *no* responses than at time 1. This supports the need for the survey to include an option stating that they are unable to judge the patient's symptoms, as proposed in Section 8.1.2. Secondly, pain relief was not always accounted for when rating patients pain. This finding was observed in several instances at time 1. More recent versions of VOICES, instead of assessing the severity of pain, ask how effective the pain relief is from *completely some of the time* to *not at all*. This overcomes the problem of not taking into account pain relief and supports VOICES approach to assessing pain. The change in symptom focus (e.g. pain to oedema) observed in a few instances could be attributed to the focused nature of the research because only pain, anxiety and depression were assessed. However, proxies should be encouraged to include symptoms that they think are important, so they are not expressed within other rating scales. For example, itchiness was a distressing symptom for one participant and his wife, yet there is no way of expressing it in VOICES. Additional spaces should be included in the questionnaire so that this type of additional information can be assessed. The measures recommended may reduce the degree of incongruence and proxies' inclination to overstate patients' pain retrospectively (Higginson *et al.*, 1994; Hinton, 1996).

In contrast to pain, the consistency of anxiety ratings was extremely good (moderate to substantial) and could be explained in part by the saliency of events in memory associated with anxiety provoking events. As mentioned in Section 7.4.6, the level of detail recalled by

some proxies was both remarkable and consistent in both the time periods examined. Events, such as admission to hospital or hospice, symptom control and patient diagnosis were viewed as the catalyst for much of the patient's and proxy's anxiety. This may explain the finding that proxies retrospective assessments of patients' anxiety are more in agreement with the patients own rating (Hinton, 1996).

It is difficult to draw equally definitive conclusion on the matter of depression. Although the results suggest that there are substantial changes in proxies' ratings over time, there were too few cases where depression was in evidence. Consequently, the findings need to be viewed with caution. For those cases that were identified, there were low levels of agreement in the month, with more uncertainty expressed and greater numbers reporting no depression at time 2. Agreement was better during the week, with fair to moderate levels observed. As it did with pain, the patient's condition appeared to contribute to the uncertainty, particularly when the patient became confused or there was a decrease in their level of consciousness. The proxy's construction of events coincided with some changes in their ratings of both depression and pain. Changes in the way similar events were viewed is probably a function of the way in which the proxy is coping with events. In general, ratings became less frequent and severe, although some ratings did not alter. In a few instances, ratings even became more frequent and severe. The present work concurs with studies that have examined the validity of proxies retrospectively, where depression is underestimated and pain becomes polarised (Higginson *et al.*, 1994; Hinton, 1996). Future research should focus on the assessment of proxies' capacity to cope, in relation to their interpretations of events and VOICES ratings, in an effort to elucidate possible relationships. Moreover, the relationship between proxies' psychological status should be studied in connection with changes over time to find out whether their emotions play a role in the fluctuations. There was some hint of this with two proxies where they admitted that their feelings had altered the way in which they perceived events.

Another finding that was identified in the analysis of research questions two and four concerns reference periods. Although some proxies did talk about events outside the reference periods (week, month), there was no way of determining whether or not they influenced their responses. That is are they including information from outside the reference period in their estimates. Examining proxies' responses over time, it was possible to detect the likely affects of forward telescoping on their VOICES ratings. Whenever the ratings were within the reference period, the ratings were lower, thereby supporting previous research (Sudman & Bradburn, 1973; Bradburn *et al.*, 1979).

### 8.1.5 Contributions to the field of research

The research presented in this Ph.D. is both novel in its aims and unique in its approach to the study of proxies for patients at the end of life. The validity of proxies has been evaluated in various areas of health care including palliative care, epidemiology, decision making, and quality of life. Other investigators have published reviews in specific areas, for example for patients with chronic diseases (Sprangers & Aaronson, 1992; Lobchuk & Degner, 2002) and epidemiology research (Nelson *et al.*, 1990). However, a review of the literature of proxies for patients with terminal illness had not been undertaken before. Since significant others are often the source of information for patients, a systematic review of the current evidence was performed. Clear comparisons were drawn across the literature with regard to the type of information (e.g. objective and subjective), characteristics associated with the proxy (e.g. caregiver burden) and patient (e.g. impairment), the assessment itself (e.g. detail required), and their affects on the accuracy of proxies' responses. The review adds to the field, by broadly analysing the use and accuracy of proxies in general. It went further than others by locating the work within the wider context of social and cognitive psychology, survey methodology and palliative care literature. As a consequence, it provides a greater understanding of the possible reasons for incongruence in proxy reporting. For investigators in the area of palliative care research and audit, the review is informative because it highlights areas where proxy reports are likely to be more accurate than others, for instance when reporting on more concrete, observable aspects of the patients experiences. It also provides factors to consider when using proxy derived data, such as caregiver burden and the amount of contact between the patient and proxy.

Using the review as basis, this Ph.D. explored areas where discrepancies were known to exist and to bridge some of the gaps in the knowledge. For instance, it is known that caregiver strain is associated with levels of agreement, but little is known about how emotions influence judgments within this context. Accordingly, the reliability of responses during the bereavement period is not well understood. The information proxies use to form a judgment about patients' symptoms are particularly important for assessing the validity of proxies' responses. Research question two identified a variety of cues and knowledge upon which proxies draw. Apart from showing that proxies attempt to extract information on the actual experiences, these findings provide researchers with a means to design measures so that they can better assess symptoms. This could include a number of the overt manifestations such as medication usage and their facial expressions in relation to pain. This Ph.D. represents the first investigation of these issues.

The present research is grounded in well-established bodies of literature and the findings have theoretical as well as practical significance. Theories that have been utilised within



the context of health care and palliative care to expound coping, have been equally applicable in the context of the present work. For example, Folkman and Lazarus' (1984) theory of stress explains changes in proxies' responses over time in terms of a reappraisal of events. In addition to this, theories from other disciplines have been applied. In many cases, this has not been done before. As an example, the cognitive psychology literature on emotions and memory was used to raise awareness of issues likely to affect proxies' responses during bereavement. Bower's (1981; 1987) mood congruency literature was drawn upon to explain possible biases in proxies' responses as a result of their own feelings of anxiety and depression. One of the criticisms of this theory is the automatic connections between feelings and cognition and the passive role of cognitions (Swartz & Clore, 1989; Wyer *et al.*, 1999). Instead, Swartz and Clore (1988) propose that emotions can play an active part in the processing of information. For example, in chapter 6, affect based judgments were identified in proxies' verbalisations. Although the investigations were not a test of what Swartz and Clore (1988) call affect-based judgments, or Bower's (1981; 1987) mood congruency theory, the findings do indicate that proxies' own feelings may act as information.

Another difficulty for Bower's (1981; 1987) mood congruency theory is that not all studies have found the effects described therein, which suggests that other processes may intercede (Forgas, 1991; Mogg *et al.*, 1987; Mayer *et al.*, 1995). One possibility is that regulatory processes, where individuals try to improve their mood by thinking of more positive information may mediate the influence of mood on memory (Isen *et al.*, 1978). For bereaved individuals, this could be a method of coping with their loss. For instance, several proxies recalled happier occasions when asked about their relatives' symptoms. This was exemplified in the theme reminiscence in Chapter 7. Another intervening process may be motivation. Wegner and Petty (1997) found that when individuals are motivated to be accurate and aware of the potential of their feelings to bias responses, they tend to counteract the effects, often over correcting. The importance of gathering information from proxies was stressed to the participants in this study. It is possible to surmise that some proxies may have taken the influence of their own feelings into consideration, in order to be accurate in their judgments. Two proxies actually acknowledged that their feelings had biased their judgments at time 1. This was taken into account in their evaluations at time 2. Therefore, it is not inconceivable that others may have done so too. Thus, the relationship appears to be far more complex than envisaged in Bower's theory (1981; 1987).

Caution also needs to be exercised when comparing studies conducted in experimental settings, with those in the context of bereavement. Firstly, under experimental conditions, emotions are evoked by the presentation of material. These moods are transitory and do not reflect the complex emotion of grief. Secondly, when interviewing a bereaved relative, the

period of time between the experience and the recollection is much longer than under strict experimental conditions. Finally, the information sought in palliative care surveys is connected with other information and experiences, unlike recall in experimental conditions. Studies in naturalist settings, such as those from eyewitness testimony, also differ from the emotions experienced during bereavement. Consequently, theories based on these types of investigations may not always generalise to the palliative care setting.

Research methods and theories on the cognitive aspects of surveys (Oksenberg & Cannel, 1977; Tourangeau, 1984; Tourangeau & Rasinski, 1988; Willis *et al.*, 1991) provided a foundation upon which to identify and elucidate the levels of comprehension, memory and judgments used by individuals when answering questions. This is the first time that it has been applied to palliative care surveys and it offers a way to gather a wealth of information and understanding regarding proxies' responses. Using it, it was possible to identify response difficulties with the instrument itself and to effectively evaluate patients' symptoms at the end of life. The identification of strategies used in making judgments, as well as content based information raised awareness of the basis for responses and mental operations that could contribute to error. The method complements quantitative methods used in the validity studies, by adding to and enriching some of the findings. Although the research focus was on VOICES, the work and recommendations may contribute to the development of new and existing measures whenever proxies are used. The appropriateness of the methods to palliative care surveys will be discussed in the methods and design section (8.3.2).

## 8.2 Future research

Throughout the thesis and Section 8.1, future research questions and avenues of investigation have been presented. Additional work is also required to assess the psychometric properties of the VOICES questionnaire. The reliability of the questionnaire could not be assessed in this study, as the numbers were small. Any changes as a result of the recommendations also need to be evaluated to see whether or not they improve the accuracy of responses and the degree of congruency between the patient and proxy. The research presented in this Ph.D. is based primarily on the interpretations of participants who are of caucasian western European ethnicity. A similar approach with individuals from other cultural and ethnic backgrounds could be useful in developing appropriate terminology for culturally specific VOICES questionnaires. For example, the hospice in south east London that was used for the present investigations serves a population covering a wide ethnic diversity.

The Caregiver Strain Inventory (Robinson, 1983) was the instrument chosen to assess proxies' perceptions of the burden of caring for their dying relative. The instrument primarily focuses on tangible domains such as employment, finances, physical, social and time. Critics

have argued that tools such as the CSI take a narrow view of the dynamic and multidimensional nature of caregiving (Nolan *et al.*, 1994; Grant *et al.*, 1998). As Grant *et al.*, (1998) points out, “many of the central tasks which create real challenges for caregivers are those tied to preserving the dignity and self-esteem of dependant relatives” (p.60). Attempts to conceal the caregiver’s role are a common stress for caregivers (Nolan *et al.*, 1994). Indeed during this study several anecdotes support this observation. Yet this was not captured using the CSI. Future research may benefit from using methods better suited to assessing the multidimensional nature of caregiving. This could include both the negative and positive aspects of providing care. These methods would allow researchers to address other questions, such as, does the strengthening of family ties improve communication leading to better agreement, or whether particular facets of caring influence proxies’ perceptions of patients’ symptoms.

Anxiety and depression were the emotions chosen in the present investigations because it is known that memory and judgments can be biased by these emotions and that they are common in bereavement. The Beck Depression Inventory was employed, as this has been used in other bereavement studies (Stroebe *et al.*, 1985). However, the levels of depression detected were small. Whether this was because the levels were truly low or because of the measure itself is not known. Grief specific measures, such as the Revised Grief Inventory (Sanders *et al.*, 1985; 1991) or Revised Texas Inventory of Grief (Faschingbauer, 1981), might have been better able to detect depression in bereavement. The study also used the State Anxiety Inventory, which is not usually used to detect anxiety in bereaved relatives even though levels above the norm for adult populations were identified. Thus, future research might be improved by enlisting grief specific measures as other emotions, apart from depression and anxiety, could be identified and compared with proxies’ reports. Support for this research comes from the various emotions expressed in proxies’ responses (e.g. anger, guilt).

Given the paucity of literature on the validity of proxies’ accounts retrospectively one suggestion would be to assess proxies’ responses, using VOICES over time in a larger sample. This could be integrated into the design of an existing survey and would enable researchers to identify other factors that might influence proxies’ reports, such as patient and proxy characteristics. Furthermore, the reliability of the VOICES questionnaire could be ascertained.

In the present investigations there was no assessment from the patients to compare with proxies’ accounts. This type of prospective/retrospective design is difficult to conduct due to recruitment and attrition problems. One solution would be to compare proxies’ retrospective accounts with those of the nursing care records where patients have assessed their own symptoms before death. This would negate the need to interview patients and proxies before death. Proxies’ reports could then be compared to these records over time to establish the best time

to interview retrospectively. Moreover, including coping measures into the design would provide valuable insights into proxies' perceptions of events and coping during bereavement.

## 8.3 Limitations

### 8.3.1 Sample size

The representativeness of the sample needs consideration when interpreting the findings presented in this thesis. All the patients were known to the hospice services and in the majority of cases the proxy was the main caregiver. This means that the amount of contact and interaction they had with the patient was high and most were actively involved in symptom management. In practise, retrospective surveys include data from a variety of sources, and not just proxies selected through the hospice services. In some instances, proxies may not have access to all of the information requested. For example, in this study many *don't know* responses and inconsistencies in pain ratings occurred when the patient was under the care of the hospice during the last week of life.

The sample might not be representative of deaths in general in regard to patients' symptoms. The numbers with pain were higher (83% in the week and 86% in the last month of life) than levels reported in other studies (Brescia *et al.*, 1992; Vainio & Auvinen, 1996). Accordingly, proxies' reports of patients' anxiety and depression were more prevalent than has been reported elsewhere (McDonnell, 1989; Hardman *et al.*, 1989; Coyle *et al.*, 1990; Addington-Hall *et al.*, 1992). The numbers of patients admitted to a hospice in the last stages of their illness indicate that their carers were experiencing problems managing at home. The guilt associated with this might have biased their views of patients' symptoms. Several proxies had received follow-up care and some had attended a support group. This might have alleviated some of their distress since palliative care has been shown to reduce psychological distress (McCorkle, 1998, cited in Lev & McCorkle, 1998). This could account for the lower levels of depression found in the bereaved relatives interviewed. Alternatively, proxies who are more depressed might have chosen not to participate in the study, especially given the topics discussed. Therefore, the sample may differ from those who are not receiving such care or on characteristics that were not compared (e.g. depression, anger).

The response rate was comparatively low (35%) in these investigations, probably due to the sensitivity of the topics discussed and the opt-in approach chosen for recruitment. A number (20%) of potential participants could not be contacted because they lived outside the area. While a further 6% had no relative or contact address. These exclusions highlight a problem inherent in using proxies. For example, in the follow-up study four proxies had moved away and could not be contacted at time 2. This adds weight to Cartwright *et al.*'s (1973) findings, where a greater proportion of proxies could not be contacted at nine months.

Despite these problems no differences were found between those who took part and those who did not. Therefore, the sample is representative of deaths known to hospice services.

### 8.3.2 Methods and design

Since there is no assessment from either the patient or the proxy during the terminal stage of the disease, it is not possible to infer what their responses would have been prior to the death, or how these might have changed retrospectively. This limits the ability of the research to assess the validity of the retrospective approach. It is also not possible to say whether the proxy's account is a valid representation of the patient's experiences. Moreover, the investigation will not reveal whether or not an account given at one interview is more or less valid than an account given at another. Given the constraints of the timeline and the resources afforded by a Ph.D, it is not possible to execute a broader investigation. Moreover, difficulties associated with recruitment and attrition, which is frequently encountered in palliative care research (McWhinney *et al.*, 1994), could hamper attempts to assess patients and follow up proxies at varying intervals in bereavement.

The methods and framework of the analysis in the present work are located within an information processing paradigm, where the emphasis is on the mental processing that underlies individuals' responses. The validity of the approach has been ascertained in various areas that have examined attitudes and behaviour (Bolton, 1991; 1993; DeMaio *et al.*, 1993; Bickart & Felcher, 1994). However, neither the verbal protocol analysis nor the deductive analysis framework has been used in the context of palliative care before. Using a framework based on models of question answering (e.g. Oksenberg & Cannell, 1977; Tourangeau, 1984; Tourangeau & Rasinski, 1988; Willis *et al.*, 1991) it was possible to find evidence to support the stages proposed. However, the stages posited in models such as Willis *et al.* (1991) represent the optimal response sequences. In reality, numerous factors can influence how individuals actually respond. One such factor is motivation, which is likely to influence the amount of effort the individual puts into thinking about the question ((Di Jkstra & van der Zouwen, 1987). In this study, the relevance of the information, in terms of personal significance, is likely to motivate proxies to think more carefully about their answers. However, the cognitive effort expended throughout the interview, such as the difficulty experienced in answering some of the questions and the emotive nature of the inquiry, may have lead to satisficing. As noted in the review in Chapter 2, this is where an answer that appears satisfactory is chosen and partial or no integration and retrieval of information is performed (Krosnick, 1991).

In Chapter 3 the assumptions regarding verbal protocol analysis are detailed, namely that individuals have conscious access to relevant cognitive processes (Ericsson and Simon, 1980, 1984). On the whole, proxies were able to provide valuable insights into their thinking.

When answering questions about symptom severity, however, there were several instances where proxies were unable to verbalise their thinking, even though they provided a response. One possible reason for this was that the proxies did not have access to some of the cognitive processes that underlay their responses. This could include processing that occurs at the pre-cognitive stage or that is based on imagery rather than language (Glass & Arnkoff, 1982). Another possible reason is that proxies avoided thinking too much about the questions in order to avoid distress, which in turn was reflected in their verbalisations. During the interviews, few actually stated that they did not want to think about the question. In two particular cases though, proxies referred back to past answers when responding to present questions. In both instances when they were asked to verbalise their thinking in relation to questions about the patient's depression, they stated that it was the same as their anxiety response. Whether this was to avoid thinking about the questions or because they could not differentiate between anxiety and depression is not known. Other ways of avoiding thinking thoroughly about the question might have been rapid responses. Clearly, more work is required in order to better determine the validity of the approach in the field of palliative care research.

Given that the emphasis is very much on the individual, the methods chosen fail to account for context effects, such as the interaction between the individual and the interviewer. Yet this can impact on the responses given. As Sudman *et al.* (1996) points out the interview itself should be seen as a social interaction, where assumptions are made about the communications. In normal conversation, utterances are interpreted within a context. Similarly, when an interviewer asks questions during an interview, the individual interprets the questions in the light of features such as the interviewer's behaviour, previous questions, and response alternatives (Sudman *et al.*, 1996). For example, the way the questions are asked, the interviewer's interview style and responses, prompts and the topic can all have a bearing on the type and amount of information expressed. Since VOICES provided the initial motivation for this Ph.D. the questions in this study were very specific to this measure and were structured around it. Less structured approaches might have yielded different results or explained aspects that the present investigations could not. For instance, it might have elucidated some of the changes that were observed over time. Furthermore, the interviewer may have taken a less passive role.

It is important to acknowledge that all research is value-laden as it is designed and conducted. It can also be interpreted in particular ways depending on the researcher's goals and background. The methods chosen for these investigations are primarily located within cognitive psychology with an emphasis on the individual and their understanding of the phenomenon of interest. Therefore, the experiences of the investigator and their perceptions as an individual who has experienced bereavement, worked as a Registered Nurse and studied psy-

chology, are likely to influence how the proxies' responses were interpreted. For example, knowledge of psychological theories of coping was used in interpreting changes in proxies' responses over time. Although these theories did not guide the thematic analysis used in addressing research question four, they were explanatory following the analysis process.

## **8.4 Interviewing bereaved relatives**

Given the sensitive nature of the interviews and the timing I was acutely aware of the possibility of distressing participants. During the interviews many participants were tearful as they recalled events. Expressions of anger, guilt, shock and sadness were also evident. At times when participants became distressed the interview was stopped and they were asked if they wanted to proceed. All were insistent on continuing so time was given as necessary. In two instances the interviews were shortened by excluding some of the questions about the VOICES symptom descriptors and prompting was kept to a minimum. This was based on my feelings regarding the participant's level of comfort with the interview. In almost all cases the bereaved relatives felt that the interviews were beneficial in helping others and/or in enabling them to talk about their feelings. For three participants it was the first time they had actually talked about events for fear of upsetting other family members. So they were appreciative of the opportunity to share their experiences with someone from outside the family. A good rapport was built with many of the participants during the interviews. On average each visit lasted between two to three hours. This rapport was important because it facilitated the open expression of both negative and positive experiences. To maintain the flow of the interviews and build rapport the interviews did not always adhere to the layout of the PSCI. For example, pain relief was often the first point proxies made when talking about pain. Flexibility during the interviews enabled participants to fully express themselves and their needs, providing support and enhancing the quality of the data collected.

## **8.5 Conclusion**

The findings of this Ph.D. have significant implications for the retrospective approach in palliative care. If investigators ignore the potential biases introduced by using proxies (significant others and health professionals), then estimates of some symptoms will be inaccurate. The potential consequence is that any evaluation of the adequacy of services and care to patients at the end of life will be unreliable. The findings presented here highlight where the potential biases and inaccuracies lie and presents a number of recommendations to compensate for them. How and to what extent these measures are implemented depends on the level of accuracy that is required. Nonetheless, it is important to note that proxies' judgments regarding the patients' symptoms of pain, anxiety and depression indicated that they have

some validity because they were basing their judgments on observable events, and not just reflecting on their own distress. Thus, pursuing ways to improve the validity of their responses, rather than discounting them is a worthwhile objective for future research.



## References

- Aaronson, N.K. (1991) Methodological Issues in Assessing the Quality of Life of Cancer Patients. *Cancer*, 67, 844-850.
- Addington-Hall, J., Walker, L., Jones, C., Karlsen, S. & McCarthy M. (1998) A randomised controlled trial of postal versus interviewer administration of a questionnaire measuring satisfaction with, and use of, services received in the year before death. *Journal of Epidemiology & Community Health*, 52(12), 802-807,
- Addington-Hall, J., Lay, M., Altmann, D. & McCarthy, M. (1995) Symptom control, communication with health professionals, and hospital care of stroke patients in the last year of life as reported by surviving family, friends, and officials. *Stroke*, 26, 2242-2248.
- Addington-Hall, J.M., MacDonald, L.D., Anderson, H.R., Chamberlain, J., Freeling, P., Bland, J.M. & Raftery, J. (1992) Randomised controlled trial of effects of coordinating care for terminally ill cancer patients. *British Medical Journal*, 305(6865), 1317-1322.
- Addington-Hall, J.M., MacDonald, L.D., Anderson, H.R., & Freeling, P. (1991) Dying from cancer: the views of bereaved family and friends about the experiences of terminally ill patients. *Palliative Care*, 5, 207-214.
- Addington-Hall, J. & McCarthy, M. (1995a) Regional Study of Care for the Dying: methods and sample characteristics. *Palliative Medicine*, 9(12), 27-35.
- Addington-Hall, J.M. & McCarthy, M. (1995b) Dying from cancer: results of a national population-based investigation. *Palliative Medicine*, 9(4), 295-305.
- Ahles, T.A., Blanchard, E.B. & Ruckdeschel, J.C. (1983) The multidimensional nature of cancer-related pain. *Pain*, 17, 277-288.
- Ahles, T.A. & Martin, J. B. (1992) Cancer pain: A multidimensional perspective. *Hospice Journal*, 8(1-2), 25-48.
- Ahmedzai, S., Morton, A., Reid, J. & Stevenson, R. (1988) Quality of death from lung cancer: patients' reports and relatives' retrospective opinions. In: M.Watson, S.Greer & C.Thomas, (Eds.) *Psychosocial Oncology*. Oxford: Pergamon Press. pp.187-192.
- Aikman, P.J., Thiel, E.C., Martin, D.K. & Singer, P.A. (1999) Proxy, Health, and Personal Care Preferences: Implications for End-of-Life Care, *Cambridge Quarterly of Healthcare Ethics*, 8, 200-210.
- Altheide, D.L (1987) Ethnographic content analysis *Qualitative Sociology*, 10(1), 65-67.
- American Psychiatric Association. (1994) *Diagnostic and Statistical Manual of Mental Disorders* (4<sup>th</sup> Ed.). Washington DC: Author.
- Anonymous. (2000) Survey into the health problems of elderly people: a comparison of self-report with proxy information. *International Journal of Epidemiology*, 29(4), 684-697.
- Armstrong, A.M., MacDonald, A., Booth, I.W., Platts. R.G., Knibb, R.C & Booth, D.A. (2000) Errors in memory for dietary intake and their reduction. *Applied Cognitive Psychology*, 14, 183-191.
- Au, E., Loprinzi, C.L., Dhodapkar, M., Nelson, T., Novotny, P., Hammack, J. & O'Fallon, J. (1994) Regular use of verbal pain scale improves the understanding of oncology inpatient pain intensity. *Journal of Clinical Oncology*, 12(12), 2751-2755.
- Baddley, A. (1990) *Human Memory: Theory and Practice*, London: Erlbaum Associates.
- Baider, L. & Sarell, M. (1984) Couples in crisis: Patient-spouse differences in perception of interaction patterns and illness situation. *Family Therapy* 9, 115-122.

- Barbour, R.R. (1999) The case for combining qualitative and quantitative approaches in health services research. *Journal of Health Services Research Policy*, 4, 39-43.
- Barclay, C.R. & DeCooke, P.A. (1988) Ordinary everyday memories ... some of the things of which selves are made. In: U. Neisser & E. Winograd (Eds.) *Remembering reconsidered: ecological and traditional approaches to the study of memory*. New York: Cambridge University Press. pp.91-125.
- Barracough, J. (1997) Depression, anxiety, and confusion. *British Medical Journal*, 315, 1365-1368.
- Bartlett, F.C. (1932) *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press. pp.309.
- Bartrop, R.W., Hancock, K., Craig, A. & Porritt, D.W. (1992) Psychological toxicity of bereavement: six months after the event. *Australian Psychologist*, 27, 192-196.
- Bassett, S.S., Magaziner, J. & Hebel, J.R. (1990) Reliability of proxy response on mental health indices for aged, community-dwelling women. *Psychology & Aging*, 5, 127-132.
- Beck, A.T. (1963) Thinking and depression: I. Idiosyncratic content and cognitive distortions. *Archives of General Psychiatry*, 24, 324-333.
- Beck, A.T., Emery, G. & Greenberg, R.L. (1985) *Anxiety Disorders and Phobias: A Cognitive Perspective*. New York: Basic Books.
- Beck, A.T., Rush, A.J., Shaw, B.F. & Emery, G. (1979) *Cognitive Therapy of Depression*. Chichester: Wiley.
- Beck, A.T., Steer, R.A. & Brown, G.K. (1996) *Manual for the Beck Depression Inventory BDI-II*. San Antonio, TX: The Psychological Corporation.
- Beck, A.T., Steer, R.A. & Garbin, M.G. (1988) Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8, 77-100.
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J. & Erbaugh, J. (1961) An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Belson, W.A. (1981) *The Design and Understanding of Survey Questions*. London: Gower.
- Bickart, B. & Felcher, E.M. (1996) Expanding and Enhancing the Use of Verbal Protocols in Survey Research. In: N. Schwarz & S. Sudman. (Eds.) *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Methodology*. San Francisco: Jossey-Bass Publishers. pp.115-142.
- Bickart, B., Menon, G., Schwarz, N. & Blair, J. (1994) The use of Anchoring Strategies in Constructing Proxy Reports of Attitudes. *International Journal of Public Opinion Research*, 6, 375-379.
- Biley, F. (1989) Nurses' perception of stress in preoperative surgical patients. *Journal of Advanced Nursing*, 14(7), 575-581.
- Blair, E. & Burton, S. 1987, Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research*, 14, 280-288.
- Blair, J., Menon, G. & Bickart, B. (1991) Measurement effects in self versus proxy response to survey questions: An information processing perspective. In: P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz & S. Sudman.(Eds.) *Measurement Errors in Surveys*. New York: Wiley. pp.145-166.
- Blaney, P.H. (1986) Affect and memory: A review. *Psychological Bulletin*, 99, 229-246.

- Blazeby, J.M., Williams, M.H., Alderson, D. & Farndon, J.R. (1995) Observer variation in assessment of quality of life in patients with oesophageal cancer. *British Journal of Surgery*, 82, 1200-1203.
- Bolton, R.N. & Bronkhorst, T.M. (1991) Quantitative analyses of depth interviews. *Psychology and Marketing*, 8, 275-297.
- Bolton, R.N. (1993) Pretesting questionnaires: content analyses of respondents' concurrent verbal protocols. *Marketing Science*, 12, 280-303.
- Bonanno, G.A., Keltner, D., Holen, A. & Horowitz, M.J. (1995) When avoiding unpleasant emotion might not be such a bad thing: Verbal-autonomic response dissociation and midlife conjugal bereavement. *Journal of Personality and Social Psychology*, 46, 975-989.
- Bower, G.H. (1981) Mood and Memory. *American Psychologist*, 36, 129-148.
- Bower, G.H. (1987) Commentary on mood and memory. *Behaviour Research and Therapy*, 25, 443-455.
- Bower, G.H. (1992) How might emotion affect learning? In: S-A. Christianson (Ed.) *Handbook of emotion and memory*. Hillsdale, NJ: Erlbaum. pp.3-31.
- Bower, G., Black, J., Turner, T. (1979) Scripts in memory for text. *Cognitive Psychology*, 11, 177-120.
- Bowlby, J. (1969) *Attachment and Loss, Volume 1. Attachment*. London: The Hogarth Press.
- Bowlby, J. (1980) *Attachment and Loss, Volume 3. Loss: Sadness and Depression*. London: The Hogarth Press.
- Bowlby, J & Parkes, C.M. (1970) Separation and loss within the family. In: E.J. Anthony & C. Koupornik (Eds.) *The Child and His Family*. London: Wiley. pp.197-216.
- Bradburn, N.M., Rips, L.J. & Shevell, S.K. (1987) Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys. *Science*, 236, 157-161.
- Bradburn, N.M. & Sudman, S. (1979) *Improving Interview Method and Questionnaire Design: Response Effects to Threatening Questions in Survey Research*. San Francisco: Jossey-Bass.
- Bradley, B.P., Mogg, K. & Lee, S.C. (1997) Attentional biases for negative information in induced and naturally occurring dysphoria. *Behaviour Research & Therapy*, 35(10), 911-927.
- Breitbart, W. & Holland, J. (1988) Psychosocial aspects of head and neck cancer. *Seminars in Oncology*, 15(1), 61-69.
- Breitbart, W. & Jacobsen, P.B. (1996) Psychiatric Symptom Management in Terminal Care. *Clinics in Geriatric Medicine*, 12(2), 329-347.
- Breitbart, W., Bruera, E., Chochinov, H. & Lynch, M. (1995) Neuropsychiatric syndromes and psychological symptoms in patients with advanced cancer. *Journal of Pain and Symptom Management*. 10(2), 131-141.
- Brescia, F.J., Portney, R.K., Ryan, M., Krasnoff, L. & Gray, G. (1992) Pain, opioid use, and survival in hospitalised patients with advanced cancer. *Journal of Clinical Oncology*, 10, 149-155.
- Brewer, W. (1986) What is autobiographical memory? In: D.C. Rubin (Ed.) *Autobiographical memory*. New York: Cambridge University Press. pp.25-50.

- Briggs, M. & Closs, J.S. (1999) A descriptive study of the use of visual analogue scales and verbal rating scales for the assessment of postoperative pain in orthopedic patients. *Journal of Pain and Symptom Management*, 18(6), 438- 446.
- British Psychological Society. (2000) Ethical Principles and Guidelines for Psychologists. <http://www.bps.org.uk/about/rules5.cfm>. [Accessed 24/01/2000].
- Broadbent, D. & Broadbent, M. (1988) Anxiety and attentional bias "State and trait". *Cognition and Emotion*, 2, 165-183.
- Brown, N.R., Shevell, S.K. & L.J. Rips. (1986) Public memories and their personal content. In: D.C. Rubin (Ed.) *Autobiographical Memory*. Cambridge: Cambridge University Press. pp.137-158.
- Bruera, E., Kuehn, N., Miller, M., Selmser, P. & Macmillianm K. (1991) the Edmonton Symptom Assessment System (ESAS): a simple method for the assessment of palliative care patients. *Journal of Palliative Care*, 7, 6-9.
- Brunelli, C., Costantini, M., Di Giulio, P., Gallucci, M., Fusco, F., Miccinesi, G., Paci, E., Peruselli, C., Morino, P., Piazza, M., Tamburini, M. & Toscani, F. (1998) Quality-of-Life Evaluation: When Do Terminal Cancer Patients and Health-care Providers Agree? *Journal of Pain and Symptom Management*, 15(3), 151-158.
- Brugha, T.S. (1993) Depression in the terminally ill. *British Journal of Hospital Medicine*, 50, 175-181.
- Bryant, R.A. (1993) Memory for pain and affect in chronic pain patients. *Pain*, 54, 347-351.
- Burke, M. & Matthews, A. (1992) Autobiographical memory and clinical anxiety. *Cognition and Emotion*, 6, 23-35.
- Burton, S. & Blair, J. (1991) Task conditions, response formulation processes, and response accuracy for behavioral frequency questions in surveys. *Public Opinion Quarterly*, 55, 50-79.
- Cacioppo, J.T., von Hippel, W. & Ernst, J.M. (1997) Mapping Cognitive Structures and Processes Through Verbal Content: The Thought-Listing Technique. *Journal of Consulting and Clinical Psychology*, 65, 928-940.
- Cannell, C.F., Blixit, S., & Oksenberg, C. (1992) *Characteristics of Questions That Cause Problems for Interviewers or Respondents*. Ann Arbor MI: Survey Research Center.
- Capuzzo, M., Grasselli, C., Carrer, S., Gritti, G. & Alvisi, R. (2000) Quality of life before intensive care admission: agreement between patient and relative assessment. *Intensive Care Medicine*, 26(9), 1288-1295.
- Carson, H. (1932) The learning and retention of pleasant and unpleasant activities. *Archives of Psychology*, 134, 1-96.
- Cartwright, A. & Seale, C. (1990) *The Natural History of the Survey: An account of the methodological issues encountered in a study of life before death*. London: King's Fund.
- Cartwright, A., Hockey, L. & Anderson, J.L. (1973) *Life before death*. London: Routledge and Kegan Paul.
- Cella, D.F., Jacobson, P.B., Orav, E.J., Holland, J.C., Silberfarb, P.M. & Rafla, S. (1987) A brief POMS measure of distress for cancer patients. *Journal of Chronic Disease*, 40(10) 939-942.
- Chamberlain, K. (1999) Using Grounded Theory in Health Psychology. In: M. Murray & K. Chamberlain (Eds.) *Qualitative Health Psychology: Theories and Methods*. London: Sage Publications. pp.183-201.

- Charmaz, K. (1995) Grounded theory. In: J. Smith, R. Harre & L. Van Langenhove (Eds.) *Rethinking Methods in Psychology*. London: Sage Publications. pp.27-49.
- Chochinov, H.M. (2000) Psychiatry and Terminal Illness. *Canadian Journal of Psychiatry*, 45, 143-150.
- Christianson, S-A (1992) Emotional stress and eyewitness memory: a critical review. *Psychological Bulletin*, 112, 284-309.
- Christianson, S-A. & Loftus, E.F. (1987) Memory for traumatic events. *Applied Cognitive Psychology*, 1, 225-239.
- Cicchetti, D.V. (1976) Assessing Inter-Rater Reliability for Rating Scales: Resolving some Basic Issues. *British Journal of Psychiatry*, 129, 452-456.
- Claridge, B.R. & Massagli, M.P. (1989) The Use of Female Proxies in Common Symptom Reporting. *Medical Care*, 27(4), 352-365.
- Clark, H.H. & Schober, M.F. (1994) Asking Questions and Influencing Answers. In: J.M. Tanur (Ed.) *Questions about Questions: Inquiries into the Cognitive Bases of Surveys*. New York: Russell Sage Foundation. pp.15-49.
- Clark, L.A. (1989) The anxiety and depressive disorders: Descriptive psychopathology and differential diagnosis. In: P.C. Kendall & D. Watson (Eds.) *Anxiety and depression: Distinctive and overlapping features*. San Diego, CA: Academic Press. pp.83-129.
- Clayton, P.J. (1990) Bereavement and depression. *Journal of Clinical Psychiatry*, 51, 34-38.
- Cleeland, C.S., Gonin, R., Hatfield, A.K., Edmonson, J.H., Blum, R.H., Stewart, J.A. & Pandya, K.J. (1994) Pain and its treatment in outpatients with metastatic cancer. *New England Journal of Medicine*, 330, 592-596.
- Clipp, E.C. & George, L.K. (1992) Patients with cancer and their spouse caregivers. Perceptions of the illness experience. *Cancer*, 69, 1074-1079.
- Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Cohen, J. (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70, 213-220.
- Cohen, J. (1988) *Statistical power analysis for the Behavioural Sciences*. Hillsdale, NJ: Erlbaum Associates.
- Colt, J.S., Engel, L.S., Keifer, M.C., Thompson, M.L. & Zahm, S.H. (2001) Comparability of data obtained from migrant farm workers and their spouses on occupational history. *American Journal of Industrial Medicine*, 40(5), 523-530.
- Conill, C., Verger, E., Henriquez, I., Saiz, N., Espier, M., Lugo, F. & Garrigos, A. (1997) Symptom prevalence in the last week of life. *Journal of Pain and Symptom Management*, 14(6), 328-331.
- Conway, M.A. (1992) A structural model of autobiographical memory. In: M.A. Conway, D.C. Rubin, H. Spinnler & W.A. Wagenaar (Eds.) *Theoretical perspectives on autobiographical memory*. Dordrecht: Kluwer Academic Publishers. pp.167-193.
- Conway, M.A. (1992) A structural model of autobiographical memory. In: M.A. Conway, D.C. Rubin, H. Spinnler, W.A. Wagenaar. *Theoretical Perspectives on Autobiographical Memory*. Amsterdam: Kluwer Academic. pp.202-221.
- Conway, M.A. (1995) *Flashbulb memories*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Conway, M.A. (1997) *Cognitive Models of Memory*. Cambridge, Massachusetts: The MIT Press.

- Coyle, N., Adelhardt, J., Foley, K.M. & Portenoy, R.K. (1990) Character of terminal illness in the advanced cancer patient: Pain and other symptoms during the last four weeks of life. *Journal of Pain and Symptom Management*, 5(2), 83-93.
- Craig, K.D. (1989) Emotional aspects of pain. In: P.D. Wall & R. Melzack (Eds.) *Textbook of Pain* (2nd Ed.) London: Churchill Livingstone. pp.220-230.
- CRD (NHS Centre for Reviews and Dissemination) (1996) Undertaking systematic reviews of research on effectiveness: CRD Guidelines for those carrying out and commissioning reviews. *CRD Report 4*. York: University of York, York Publishing Services
- Curtis, E.B., Krech, R. & Walsh, D. (1991) Common symptoms in patients with advanced cancer. *Journal of Palliative Care*, 7(2), 25-29.
- Cusick, C.P., Gerhart, K.A. & Mellick, D.C. (2000) Participant-proxy reliability in traumatic brain injury outcome research. *Journal of Head Trauma Rehabilitation*, 15(1), 739-749.
- Dar, R., Beach, C.M., Barden, P.L. & Cleeland, C.S. (1992) Cancer pain in the marital system: a study of patients and their spouses. *Journal of Pain and Symptom Management*, 7, 87-93.
- Daut, R.L. Cleeland, C.S. & Flanery, R.C. (1983). Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. *Pain*, 17, 197-210.
- Davitz, J.R. & Davitz, L.L. (1981) *Inferences of patients' pain and psychological distress: Studies of nursing behaviors*. New York: Springer.
- DeMaio, T., Ciochetto, S. & Davis, W.L. (1993) Research on the Continuing Survey of Food Intakes by Individuals. *Proceedings of the American Statistical Association Section on Survey Research Methods*, 1021-1026.
- DeMaio, T. & Rothgeb, J. (1996) Cognitive interviewing techniques: In the lab and field. In: N. Schwarz & S. Sudman. (Eds.) *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Methodology*. San Francisco: Jossey-Bass Publishers. pp.177-196.
- Derogatis, L.R., Abeloff, M.D. & McBeth, C.D. (1976) Cancer patients and their physicians in the perception of psychological symptoms. *Psychosomatics*, 17, 194-201.
- Derogatis, L.R., Morrow, G.R., Fetting, J., Penman, D., Piasetsky, S., Schmale, A.M., Henrichs, M. & Carnickle, C.L.M. (1983) The prevalence of psychiatric disorders among cancer patients. *JAMA*, 249, 751-757.
- Diamond, E.L., Jernigan, J.A., Moseley, R.A., Messina, V. & McKeown, R.A. (1989) Decision-Making Ability and Advance Directive Preferences in Nursing Home Patients and Proxies. *The Gerontologist*, 29(5), 625-626.
- Di Jkstra, W. & van der Zouwen, J. (1987) Styles of Interviewing and the Social Context of the Survey-Interview. In: H-J. Hippler, N. Swartz & S. Sudman (Eds.) *Social Information Processing and Survey Methodology*. New York: Springer-Verlag, pp.200-212.
- Donnelly, S. & Walsh, D. (1995) The symptoms of advanced cancer: identification of clinical and research priorities by assessment of prevalence and severity. *Journal of Palliative Care*, 11(1), 27-32.
- Duncan, P.W., Lai, S.M., Tyler, D., Perera, S., Rekler, D. & Studenski, D. (2002) Evaluation of Proxy Responses to the Stroke Impact Scale. *Stroke*, 33(11), 2593-2599.
- Eich, E. (1995) Searching for mood dependent memory. *Psychological Science*, 6, 67-75.

- Eisenhower, D., Mathiowetz, N.A. & Morganstein, D. (1991) Recall Error: Sources and Bias Reduction Techniques. In: P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz & S. Sudman.(Eds.) *Measurement Errors in Surveys*. New York: Wiley. pp.127-144.
- Eiser, C. & Morse, R. (2001a) Quality-of-life measures in chronic diseases of childhood. *Health Technology Assessment*, 5(4), 1-157.
- Eiser, C. & Morse R. (2001b) Can parents rate their child's health-related quality of life? Results of a systematic review. *Quality of Life Research*, 10(4), 347-357.
- Ell, K., Nishimoto, R., Mantell, J. & Hamovitch, M. (1988) Longitudinal analysis of psychological adaptation among family members of patients with cancer. *Journal of Psychosomatic Research*, 32, 429-438.
- Elliott, B.A., Elliott, T.E., Murray, D.M., Braun, BL & Johnson, K.M. (1996) Patients and family members: the role of knowledge and attitudes in cancer pain. *Journal of Pain and Symptom Management*, 12, 209-220.
- Epstein, A.M., Hall, J.A., Tognetti, J., Son, L.H. & Conant, L. Jr. (1989) Using proxies to evaluate quality of life. Can they provide valid information about patients' health status and satisfaction with medical care? *Medical Care*, 27, S91-S98.
- Ericsson, K.A. & Simon, H.A. (1980) Verbal reports as data. *Psychological Review*, 87, 215-251.
- Ericsson, K.A. & Simon, H.A. (1984) *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Ericsson, K.A. & Simon, H.A. (1993) *Protocol analysis: Verbal reports as data*. (Rev. Ed.) Cambridge, MA: MIT Press.
- Ermont, S.L., Collins, L.R. & Zywall, W.H. (1991) Methodological Note: Corroboration of Self-Reported Smoking Status Using Significant Other Reports. *Addictive Behaviors*, 16, 329-333.
- Evans, D.A. (1990) Estimated prevalence of Alzheimer's disease in the United States. *Milbank Quarterly*, 68, 267-289.
- Eysenck, M.W. (1992) *Anxiety: The cognitive perspective*. Hove, NJ: Erlbaum
- Eysenck, M.W., Mogg, K. (1992) Clinical anxiety, trait anxiety, and memory bias. In: S-A Christianson (Ed.) *The handbook of emotion and memory: Research and theory*. Hillsdale, NJ: Erlbaum. pp. 429-450
- Fainsinger, R., MacEachern, T., Hanson, J., Miller, M. & Bruera, E. (1991) Symptom control during the last week of life on a palliative care unit. *Journal of Palliative Care*, 7(1), 5-11.
- Farrell, G. (1991) How accurately do nurses perceive patients' needs? A comparison of general and psychiatric settings. *Journal of Advanced Nursing*, 16(9), 1062-1070.
- Farrow, D.C. & Samet, J.M. (1990) Comparability of information provided by elderly cancer patients and surrogates regarding health and functional status, social network, and life events. *Epidemiology*, 1, 370-376.
- Faschingbauer, T.R. (1981) *Texas revised inventory of grief manual*. Houston, TX: Honeycomb Publishing.
- Ferrell, B.R., Ferrell, B.A., Rhiner, M. & Grant, M. (1991) Family factors influencing cancer pain management. *Postgraduate Medical Journal*, 67(Suppl.2), S64-S69.
- Fewtrell, D. & O'Connor, K. (1995) *Clinical Phenomenology and Cognitive Psychology*. New York: Routledge. pp.30-31.

- Field, D., Douglas, C., Jagger, C. & Dand, P. (1995) Terminal illness: views of patients and their lay carers. *Palliative Medicine*, 9, 45-54.
- Fikree, F.F., Gray, R.H. & Shah, F. (1993) Can Men Be Trusted? A Comparison of Pregnancy Histories Reported by Husbands and Wives. *American Journal of Epidemiology*, 138(4), 237-242.
- Fishman, B., Pasteranak, S., Wallenstien, S.L., Houde, R.W., Holland, J.C. & Foley, K.M. (1987) The memorial pain assessment card: a valid instrument for the evaluation of cancer pain. *Cancer*, 60, 1151-1158.
- Folkman, S. (1997) Positive psychological states and coping with severe stress. *Social Science and Medicine*, 45, 1207-1221.
- Folkman, S., & Lazarus, R.S. (1990) Coping and emotion. In: N.L. Stein, B., Leventhal, T. Trabasso (Eds.) *Psychological and biological approaches to emotion*. Hillsdale, N.J: Erlbaum. pp.313-332.
- Forgas, J.P. (1991) *Emotion and social judgments: an introductory review*. Oxford: Pergamon Press.
- Fowler, F.J. Jr. & Cannell, C.F. (1996) Using Behavioral Coding to Identify Cognitive Problems with Survey Questions. In: N. Schwarz & S. Sudman (Eds.) *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco: Jossey-Bass. pp.15-35.
- Friedman, W.J. & deWinstanley, P.A. (1998) Changes in the subjective properties of Autobiographical memories with the passage of time. *Memory*, 6(4), 367-381.
- Gaston-Johansson, F., Albert, M., Fafan, E. & Zimmerman, L. (1990) Similarities in Pain Descriptors of Four Different Ethnic-Culture Groups. *Journal of Pain and Symptom Management*, 5(2), 94-100.
- Gilligan, S.G. & Bower, G.H. (1984) Cognitive consequences of emotional arousal. In: C. Izard, J. Kagan & R. Zajonc (Eds.) *Emotions, Cognitions and Behaviour*. New York: Cambridge University Press. pp.547-588.
- Glasser, B.G. (1978) *Theoretical Sensitivity: Advances in the Methodology of Grounded Theory*. Mill Valley, CA: Sociology Press.
- Glaser, B.G. & Strauss, A.L. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Glass, C.R. & Arnkoff, D.B. (1983) Cognitive set and level of anxiety: Effects on thinking processes in problematic situations. *Cognitive Therapy & Research*, 7(6) 529-542.
- Glickman, M. (1997) *Making palliative care better: quality improvement, multi-professional audit and standards*. Occasional Paper 12. London: National Council for Hospice and Specialist Palliative Care Services.
- Goldberg, D.P. (1972) *The detection of psychiatric illness by questionnaire*. London: Oxford University Press.
- Gotlib, I. & Cane, D. (1989) Self-report assessment of depression and anxiety. In: P.C. Kendall & D. Watson (Eds.) *Anxiety and depression: Distinctive and overlapping features*. San Diego, CA: Academic Press. pp.131-169.
- Grassi, L., Indelli, M., Maltoni, M., Falcini, F., Fabbri, L. & Indelli, R. (1996) Quality of life of homebound patients with advanced cancer: assessments by patients, family members, and oncologists. *Journal of Psychosocial Oncology*, 14, 31-45.



- Grant, G., Ramcharan, P., McGrath, M., Nolan, M. & Keady, J. (1998) Rewards and gratifications among family caregivers: towards a refined model of caring and coping. *Journal of Intellectual Disability Research*, 42(1), 58-71.
- Grbich, C. (1999) *Qualitative Research in Health: An Introduction*. London: Sage.
- Grossman, S.A., Sheidler, V.R., Swedeen, K., Mucenski, J. & Piantadosi, S. (1991) Correlation of patient and caregiver ratings of cancer pain. *Journal of Pain and Symptom Management*, 6, 53-57.
- Halabi, S., Zurayk, H., Awaida, R., Darwish, M. & Saab, B. (1992) Reliability and Validity of Self and Proxy Reporting of Morbidity Data: A Case Study from Beirut, Lebanon. *International Journal of Epidemiology*, 21(3), 607-612.
- Hammerton, M. (1976) How much is a large part? *Applied Ergonomics*, 7, 10-12.
- Hardman, A., Maguire, P. & Crowther, D. (1989) The recognition of psychiatric morbidity on a medical oncology ward. *Journal of Psychosomatic Research*, 33, 235-239.
- Hare, J., Pratt, C. & Nelson, C. (1992) Agreement between patients and their self-selected surrogates on difficult medical decisions. *Archives of Internal Medicine*, 152, 1049-1054.
- Hatch, M.C., Misra, D., Kabat, G.C. & Karyzmer, S. (1991) Proxy Respondents in Reproductive Research: A Comparison of Self- and Partner-Reported Data. *American Journal of Epidemiology*, 133(1), 826-831.
- Hays, R.D., Vickrey, B.G., Hermann, B.P., Perrine, K., Cramer, J., Meador, K., Spritzer, K. & Devinsky, O. (1995) Agreement between self-reports and proxy reports of quality of life in epilepsy patients. *Quality of Life Research*, 4, 159-168.
- Hermann, N. (1985) Retrospective information from questionnaires. Comparability of primary respondents and their next-of-kin. *American Journal of Epidemiology*, 121(6), 937-947.
- Higginson, I. (1993) *Clinical Audit in Palliative Care*. Oxford: Radcliffe Medical Press.
- Higginson, I.J. & McCarthy, M. (1993) Validity of the support team assessment schedule: do staffs' ratings reflect those made by patients or their families? *Palliative Medicine*, 7, 219-228.
- Higginson, I., Priest, P. & McCarthy, M. (1994) Are bereaved family members a valid proxy for a patient's assessment of dying? *Social Science & Medicine*, 38, 553-557.
- Higginson, I., Wade, A. & McCarthy, M. (1990) Palliative care views of patients and their families. *British Medical Journal*, 301, 277-281.
- Hinton, J. (1972) Psychiatric consultation in fatal illness. *Proceedings of the Royal Society of Medicine*, 65(11), 1035-1038.
- Hinton, J. (1979) Comparison of places and policies for terminal care. *Lancet*, 1(8106), 29-32.
- Hinton, J. (1996) How reliable are relatives' retrospective reports of terminal illness? Patients' and relatives' accounts compared. *Social Science & Medicine*, 43, 1229-126.
- Hippler, H.J., Schwarz, N. & Sudman, S. (1987) *Social information processing and survey methodology*. New York: Springer-Verlag.
- Hogan, N., Morse, J.M. & Tason, M.C. (1996) Toward an experiential theory of bereavement. *Omega*, 33(1), 43-65.
- Holing, E.V. (1986) The primary caregiver's perception of the dying trajectory. *Cancer Nursing*, 9, 29-37.

- Holmes, S. & Eburn, E. (1989) Patients' and nurses' perceptions of symptom distress in cancer. *Journal of Advanced Nursing*, 14, 840-846.
- Huber, GP & Power, DJ (1985) Retrospective reports of strategic-level managers: Guidelines for increasing their accuracy. *Strategic Management Journal*, 6, 171-180.
- Hughes, L.C. & Preski, S. (1997) Focus on quantitative methods. Using key informant methods in organizational survey research: assessing for informant bias. *Research in Nursing and Health*, 20, 81-92.
- Huskisson, E. (1974) Measurement of pain. *Lancet*, 2, 1127-1131.
- Husted, S.A. & Johnson, J.G. (1985) Oncology clients' affective states and their nurses' expectations of clients' affective states. *Cancer Nursing*, 8, 159-165.
- Ingham, J. & Portenoy, R.K. (1999) The measurement of pain and other symptoms. In: D. Doyle, G.W.C. Hanks & N. McDonald (Eds.) *Oxford Textbook of Palliative Medicine*. London: Oxford University Press. pp.203-219.
- Isen, A.M., Shalke, T.E., Clark, M. & Karp, L. (1978) Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality and Social Psychology*, 36, 1-12.
- Jabine, T., Straf, M., Tanur, J. & Tourangeau, R. (1984) *Cognitive aspects of survey methodology: Building a bridge between disciplines*. Washington DC: National Academy Press.
- Jachuck, S.J., Brierley, H., Jachuck, S. & Willcox, P.M. (1982) The effect of hypotensive drugs on the quality of life. *Journal of the Royal College of General Practitioners*, 32, 103-105.
- Jacobs, S., Hansen, F., Berkman, L., Kasl, S. & Ostfeld, A. (1989) Depressions of bereavement. *Comprehensive Psychiatry*, 30(3), 218-224.
- Jacobs, S., Hansen, F., Kasl, S., Ostfeld, A., Berkman, L. & Kim, K. (1990) Anxiety disorders in acute bereavement: Risk and risk factors, *Journal of Clinical Psychiatry*, 51, 269-274.
- Jacobs, S. & Kim, K. (1990) Psychiatric complications of bereavement. *Psychiatric Annals*, 20, 314-317.
- Jennings, B.M. & Muhlenkamp, A.F. (1981) Systematic misperception: oncology patients' self-reported affective states and their care-givers' perceptions. *Cancer Nursing*, 4, 485-489.
- Jobe, J.B. & Loftus, E.F. (1991) Cognition and survey measurement [Special issue]. *Applied Cognitive Psychology*, 5(1).
- Jobe, J.B. & Mingay, D.J. (1990) Cognitive Laboratory Approach to Designing Questionnaires for Surveys of the Elderly. *Public Health Reports*, 105, 518-524.
- Jobe, J.B. & Mingay, D.J. (1991) Cognition and survey measurement: history and overview. *Applied Cognitive Psychology* 5, 175-192.
- Jobe, J.B., Tourangeau, R. & Smith, A.F. (1993) Contributions of Survey Research to the Understanding of Memory. *Applied Cognitive Psychology*, 7, 567-584.
- Kahneman, D., Slovic, P. & Tversky, A. (1982) *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kahneman, D. & Tversky, A. (1971) Subjective Probability: A Judgment of Representativeness. *Cognitive Psychology*, 3, 450-454.
- Kahneman, D. & Tversky, A. (1973) On the Psychology of Prediction. *Psychological Review*, 80, 237-251.

- Karoly, P. & Jensen, M.P. (1987) *Multimethod assessment of chronic pain*. Oxford: Pergamon.
- Keitel, M. A., Cramer, S. H., & Zevon, M. A. (1999) Spouses of cancer patients: a review of the literature. *Journal of Counseling & Development*, 69, 163-166.
- Kelly, B., Edwards, P., Synott, R., Neil, R., Baillie, R. & Battistutta, D. (1999) Predictors of Bereavement Outcome for Family Carers of Cancer Patients. *Psycho-Oncology*, 8, 237-249.
- Kihlstrom, J.F., Cantor, N., Albright, J.S., Chew, B.R., Klein, S.B. & Niedenthal, B.M. (1988) Information processing and the study of the self. In: L. Berkowitz (Ed.) *Advances in experimental social psychology*, 17, San Diego, CA: Academic Press. pp.2-47.
- Killgore, W.D.S. (1999) The Visual Analogue Mood Scale: Can a Single Item Scale Accurately Classify Depressive Mood State. *Psychological Reports*, 85, (1238-1243).
- Kim, K. & Jacobs, S. (1991) Pathological grief and its relationship to other psychiatric disorders. *Journal of Affective Disorders*, 21, 257-263.
- Kivinen, P., Halonen, P., Eronen, M. & Nissinen, A. (1998) Self-rated health, physician-rated health and associated factors among elderly men: the Finnish cohorts of the Seven Countries Study. *Age & Aging*, 27(1), 41-47.
- Kiyak, H.A., Teri, L. & Borson, S. (1994) Physical and functional health assessment in normal aging and an Alzheimer's disease: Self-report vs family reports. *Gerontologist*, 34, 324-330.
- Klein, S.B. & Loftus, J. (1993) The Mental Representation of Trait and Autobiographical Knowledge About the Self . In: T.K. Srull, T.K. & Wyer, R.S. Jr. (Eds.) *The Mental Representation of Trait and Autobiographical Knowledge About the Self*. Hillsdale, NJ: Erlbaum. pp.1-49.
- Kolonel, L.N., Hirohata, T. & Nomura, A.M.Y. (1977) Adequacy of survey data collected from substitute respondents. *American Journal of Epidemiology*, 106, 476-484.
- Kremer, E.F. & Atkinson, J.H. (1981) Pain measurement: construct validity of the affective dimension of the McGill Pain Questionnaire with chronic benign pain patients. *Pain*, 11, 93-100.
- Kristjanson, L.J. (1989) Quality of terminal care: salient indicators identified by families. *Journal of Palliative Care*, 5, 21-30.
- Kristjanson, L.J., Nikoletti, S., Porock, D., Smith, P. & Lobchuk, M. (1998) Congruence Between Patients' and Family Caregivers Perceptions of Symptom Distress in Patients with Terminal Cancer. *Journal of Palliative Care*, 14, 24-32.
- Krosnick, J.A. (1991) Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology*, 5, 213-236.
- Kubler-Ross, E. (1969) *On Death and Dying*. New York: Macmillan.
- Kurtz, M.E., Kurtz, J.C., Given, C.C. & Given, B. (1996) Concordance of Cancer Patient and Caregiver Symptom Reports. *Cancer Practice*, 4, 185-190.
- Landis, J.R. & Koch, G.G. (1977) The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33, 159-174.
- Lang, P.J. (1985) The cognitive psychophysiology of emotion: fear and anxiety. In: A.H. Tuma & J. Maer (Eds.) *Anxiety and Anxiety Disorders*. Hillsdale, NJ: Lawrence Erlbaum. pp.131-170.

- Lawton, M.P., VanHaisma, K. & Klapper, J. (1996) Observed affect in nursing home residents with Alzheimer's disease. *Journal of Gerontology*, 51, 3-14.
- Lawton, M.P. (1997) Assessing Quality of Life in Alzheimer Disease Research. *Alzheimer Disease and Associated Disorders*, 11, 91-99.
- Layde, P.M., Beam, C.A., Broste, S.K., Connors, A.F. Jr., Desbiens, N., Lynn, J., Philips, R.S., Reding, D., Teno, J. & Vidaillet, H. (1995) Surrogates' predictions of seriously ill patients' resuscitation preferences. *Archives of Family Medicine*, 4, 518-524.
- Lazarus, R.S. (1990) Constructs of the mind in adaptation. In: N.L. Stein, B. Leventhal & T. Trabasso (Eds.) *Psychological and biological approaches to emotion*. Hillsdale, NJ: Erlbaum. pp.3-19.
- Lazarus, R.S. (1991) *Emotion and Adaptation*. Oxford: Oxford University Press
- Lazarus, R.S. & Folkman, S. (1984) *Stress, appraisal, and coping*. New York: Springer.
- Leavitt, F., Garron, D., Whisler, W. & Sheinkop, M.B. (1978) Affective and sensory dimensions of back pain. *Pain*, 4, 273-281.
- Leff, J.P. (1978) Psychiatrists' versus patients concepts of unpleasant emotions. *British Journal of Psychiatry*, 133, 306-313.
- Lerchen, M.L. & Samet, J.M. (1986) An assessment of the validity of questionnaire responses provided by surviving spouse. *American Journal of Epidemiology*, 123, 481-489.
- Lev, E.L. & McCorkle, R. (1998) Loss, Grief, and Bereavement in Family Members of Cancer Patients. *Seminars in Oncology Nursing*, 14(2), 145-151.
- Libbus, M.K. & Russell, C. (1995) Congruence of Decisions Between Patients and Their Potential Suuogates About Life-Sustaining Therapies. *IMAGE: Journal of Nursing Scholarship*, 27(2), 135-140.
- Lin, C-C. (2001) Congruity of Cancer Pain Perceptions between Taiwanese Patients and Family Caregivers: Relationship to Patients' Concerns About Reporting Pain and Using Analgesics. *Journal of Pain and Symptom Management*, 21, 18-26.
- Linton, M. (1982) Transformations of memory in everyday life. In: UN. Neisser (Ed.) *Memory observed: remembering in natural contexts*. San Francisco: WH Freeman.
- Lipton, J.A. & Marbach, J.J. (1984) Ethnicity and the pain experience. *Social Science & Medicine*, 19, 1278-1298.
- Lobchuk, M.M & Degner, L.F. (2002a) Symptom Experiences: Perceptual Accuracy Between Advanced-Stage Cancer Patients and Family Caregivers in the Home Care Setting. *Journal of Clinical Oncology*, 20(16), 3495-3507.
- Lobchuk, M.M & Degner, L.F. (2002b) Patients with Cancer and Next-of-Kin Response Comparability on Physical and Psychological Symptom Well-being. *Cancer Nursing*, 25(5), 358-374.
- Lobchuk, M.M., Kristjanson, L. (1997) Perceptions of Symptom Distress in Lung Cancer Patients: II. Behavioral Assessment by Primary Family Caregivers. *Journal of Pain and Symptom management*, 14(3), 147-156.
- Lobchuk, M.M., Kristjanson, L., Degner, L., Blood, P. & Sloan, J.A. (1997) Perceptions of symptom distress in lung cancer patients: I. Congruence between patients and primary family caregivers. *Journal of Pain and Symptom Management*, 14, 136-146.
- Loewenstain, D.A., Arguelles, S., Bravo, M., Arguelles, T., Acevedo, A. & Eisdorfer, C. (2001) Caregivers' judgments of the functional abilities of the Alzheimer's disease

- patient: a comparison of proxy reports and objective measures. *Journals of Gerontology Series B-Psychological Sciences & Social Sciences*, 56(2), 78-84.
- Long, K., Sudha, S., Mutran, E. (1998) Elder-proxy agreement concerning functional status and medical history of the older person: The impact of caregiver burden and depressive symptomology. *Journal of American Geriatrics Society*, 46, 1103-1111.
- Lynn, J., Teno, J.M., Phillips, R.S., Wu, A.W., Desbiens, N., Harrold, J., Classens, M.T., Wenger, N., Kreling, B. & Connors, A.F. Jr. (1997) Perceptions by Family Members of the Dying Experience of Older and Seriously Ill Patients. SUPPORT Investigators. Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments. *Annals of Internal Medicine*, 126, 97-106.
- MacKenzie, T.B., Robiner, W.N. & Knopman, D.S. () Differences Between Patient and Family Assessments of depression in Alzheimer's Disease. *American Journal of Psychiatry*, 149, 1174-1178.
- Madison, J.L. & Wilkie, D.J. (1995) Family members' perceptions of cancer pain: comparisons with patient sensory report and by patient psychological status. *Nursing Clinics of North America*, 30(4), 625-645.
- Magaziner, J. (1992) The use of proxy respondents in health studies of the aged. In: R.B. Wallace & R.F. Woolson. (Eds.) *The epidemiologic study of the elderly*. New York: Oxford University press. pp.120-129.
- Magaziner, J., Bassett, S.S., Hebel, J.R. & Gruber-Baldini, A. (1996) Use of proxies to measure health and functional status in epidemiologic studies of community-dwelling women aged 65 years and older. *American Journal of Epidemiology*, 143, 283-292.
- Magaziner, J., Hebel, J.R. & Warren, J.W. (1987) The use of proxy responses for aged patients in long-term care settings. *Comprehensive Gerontology*, 1, 118-121.
- Magaziner, J., Simonsick, E.M., Kashner, T.M. & Hebel, J.R. (1988) Patient-proxy response comparability on measures of patient health and functional status. *Journal of Clinical Epidemiology*, 41, 1065-1074.
- Magaziner, J., Zimmerman, S.I., Gruber-Baldini, A.L., Hebel, J.R. & Fox, K.M. (1997) proxy reporting in five areas of functional status: comparisons with self-reports and observations of performance. *American Journal of Epidemiology*, 146, 418-428.
- Mason, L. & Muhlenkamp, A.F. (1976) Patients' self-reported affective states following loss and caregivers' expectations of patients' affective states. *Rehabilitation Psychology*, 23(3), 72-76.
- Massie, M.J. (1989) Depression. In: J.C. Holland, J.H. Rowland (Eds.) *Handbook of Psychooncology: Psychology Care of the Patient with Cancer*. New York: Oxford University Press. pp.283-290.
- Mathiowetz, N.A. (1988) Forgetting events in autobiographical memory: findings from a health care survey. *Proceeding of the Section on Survey Methods Research*, Washington, DC: American Statistical Association. pp.167-172.
- Matt, G.E., Vazquez, C. & Campbell, W.K. (1992) Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review*, 12, 227-255.
- Mattimore, T.J., Wenger, N.S., Desbiens, N.A., Teno, J.M., Hamel, M.B., Liu, H., Califf, R., Connors, A.F. Jr., Lynn, J. & Oye, R.K. (1997) Surrogate and physician understanding of patients' preferences for living permanently in a nursing home. *Journal of the American Geriatrics Society*, 45, 818-824.
- Mayer, J.D., McCormick, L.J. & Strong, S.E. (1995) Mood-congruent memory and natural mood: New evidence. *Personality and Social Psychology Bulletin*, 21, 736-746.

- McCabe, S.B. & Gotlib, I.H. (1995) Selective attention and clinical depression: Performance on a deployment-of-attention task. *Journal of Abnormal Psychology*, 104(1), 241-245.
- McCarthy, M., Addington-Hall, J. & Altmann, D. (1997) The experience of dying with dementia: a retrospective study. *International Journal of Geriatric Psychiatry*, 12(3), 404-409.
- McCarthy, M., Lay, M. & Addington-Hall, J. (1996) Dying from heart disease. *Journal of the Royal College of Physicians of London*, 30, 325-328.
- McClure, M. & Willett, W.C. (1987) Misinterpretation and Misuse of the Kappa Statistic. *American Journal of Epidemiology*, 126(2), 161-169.
- McCusker, J. & Stoddard, A.M. (1984) Use of a surrogate for the Sickness Impact Profile. *Medical Care*, 22, 789-795.
- McDonnell, M.M. (1989) Patients' perceptions of their care at Our Lady's Hospice, Dublin. *Palliative Medicine*, 3, 47-53.
- McMillan, S.C. (1996) Pain and pain relief experienced by hospice patients with cancer. *Cancer Nursing*, 19(4), 298-307.
- McWhinney, I.R. (1989) Family dynamics in the achievement of symptom control. *Journal of Palliative Care*, 5, 37-39.
- McWhinney, I.R., Bass, M.J. & Donner, A (1994). Evaluation of a palliative care service: problems and pitfalls. *BMJ*, 309, 1340-1342.
- Means, B., Nigam, A., Zarrow, M., Loftus, E.F. & Donaldson, M.S. (1989) Autobiographical memory for health-related events. *Vital and Health Statistics, Series 6*(2). Washington DC: U.S. Government Printing Office. pp.1-37.
- Melzack, R. & Torgerson, W.S. (1971) On the language of pain. *Anesthesia*, 34, 50-59.
- Melzack, R. (1975) The McGill Pain Questionnaire: Major properties and scoring methods. *Pain*, 1, 277-299.
- Melzack, R. (1983) *Pain Measurement and Assessment*. New York: Raven Press.
- Melzack, R. (1987) The Short-Form McGill Pain Questionnaire, *Pain*, 30, 191-197.
- Menon, G. (1993) The effects of accessibility of information in memory of judgments of behavioral frequencies. *Journal of Consumer Research*, 20, 431-440
- Menon, G., Bickart, B., Sudman, S. & Blair, J. (1995) How Well Do You Know Your Partner? Strategies For Formulating Proxy-Reports and Their effects on Convergence to Self-Reports. *Journal of Marketing Research*, 32, 75-84.
- Miaskowski, C., Zimmer, E.F., Barrett, K.M., Dibble, S.L. & Wallhagen, M. (1997) Differences in patients' and family caregivers' perceptions of the pain experience influence patient and caregiver outcomes. *Pain*, 72, 217-226.
- Middleton, W., Moylan, A., Raphael, B., Burnett, P. & Martinek, N. (1993) An international perspective on bereavement-related concepts. *Australian and New Zealand Journal of Psychiatry*, 27, 457-463.
- Miles, M.B. & Huberman, A.M. (1994) *Qualitative Data Analysis*. 2nd Ed. London: Sage Publications.
- Mineka, S. & Sutton, S.K. (1992) Cognitive biases and emotional disorders. *Psychological Science*, 3, 65-69.
- Mogg, K., Matthews, A. & Weinman, J. (1987) Memory bias in clinical anxiety. *Journal of Abnormal Psychology*, 96, 94-98.

- Moore, M.C., Moore, E.M. & Beasley, C.H. (1970) Dietary atherosclerosis study on bereaved persons. *Journal of the American Dietary Association*, 56, 23-28.
- Morgan, D.L. (1993) Qualitative Content Analysis: A Guide to Paths Not Taken. *Qualitative Health Research*, 3(1), 112-121.
- Morris, J.N., Mor, V., Goldberg, R.J., Sherwood, S., Greer, D.S. & Hiris, J. (1986) The effect of treatment setting and patient characteristics on pain in terminal cancer patients: a report from the National Hospice Study. *Journal of Chronic Diseases*, 39, 27-35.
- Mosely, R.R. & Wolinsky, F.D. (1986) The Use of Proxies in Health Surveys: Substantive and Policy Implications. *Medical Care*, 24(6), 496-510.
- Moskowitz, J.T., Folkman, S., Collette, L., & Vittinghoff, T. (1996) Coping and mood during AIDS-related caregiving and bereavement. *Annals of Behavioral Medicine*, 18, 49-57.
- Muhlenkamp, A.F. & Joyner, J.A. (1986) Arthritis Patients' Self-Reported Affective States and their Caregivers' Perceptions. *Nursing Research*, 35(1), 24-27.
- Nekolaichuk, C.L., Bruera, E., Spachynski, K., MacEachern, T. & Maguire, T.O. (1999a) A comparison of patient and proxy symptom control assessments in advanced cancer patients. *Palliative Medicine*, 13, 311-323.
- Nekolaichuk, C.L., Maguire, T.O., Suarez-Almazor, M., Rogers, T. & Bruera, E. (1999b) Assessing the Reliability of Patient, Nurse, and Family Caregiver Symptom Ratings in Hospitalized Advanced Cancer Patients. *Journal of Clinical Oncology*, 17(11), 3621-3630.
- Nelson, L.M., Longstreth, W.T.J., Koepsell, T.D. & van Belle, G. (1990) Proxy respondents in epidemiologic research. *Epidemiologic Reviews*, 12, 71-86.
- Neter, J. & Waksberg, J. A study of response errors in expenditures data from household interviews. *Journal of the American Statistical Association*, 59, 17-55.
- Newton, T.L. & Contrada, R.J. (1992) Repressive coping and verbal-autonomic response dissociation: The influence of social context. *Journal of Personality and Social Psychology*, 62, 159-167.
- Nisbett, R.E. & Wilson, T.D. (1977) Telling more than we know: verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Nolan, M., Keady, J. & Grant, G. (1994) Developing a typology of family care: implications for nurses and other health care providers. *Journal of Advanced Nursing*, 21, 256-265.
- Novella, J.L., Jochum, C., Jolly, D., Morrone, I., Ankri, J., Bureau, F. & Blanchard, F. (2001) Agreement between patients' and proxies' reports of quality of life in Alzheimer's disease. *Quality of Life Research*, 10(5), 443-452.
- O'Brien, J. & Francis, A. (1988) The use of next-of-kin to estimate pain in cancer patients. *Pain*, 35, 171-178.
- Occupational Information Unit. (2000) *National Statistics Socio-economic Classification*. (NS-SEC). Fareham, Hampshire, Office for National Statistics.
- Okamoto, K. (2000) Agreement between the aged themselves and by their family members on physical and psychological status. *Japanese Journal of Geriatrics*, 37(5), 371-376.
- Oksenberg, L. & Cannell, C.F. (1977) Some Factors Underlying the Validity of Response in Self-Report. *International Statistical Bulletin*, 48, 325-346.
- Oksenberg, L., Cannell, C.F. & Kalton, G. (1991) New Strategies for Pretesting Survey Questions. *Journal of Official Statistics*, 7(3), 349-365.

- Ostbye, T., Tyas, S., McDowell, I & Koval, J. (1997) Reported activities of daily living: Agreement between elderly subjects with and without dementia. *Age & Aging*, 26, 99-106.
- Ouslander, J.G., Tymchuk, J. & Rahbar, B. (1989) Health Care Decisions Among Elderly Long-term Care Residents and Their Potential Proxies. *Archives of Internal Medicine*, 149, 1367-1372.
- Parkes, C.M. & Weiss, R.J. (1983) *Recovery from Bereavement*. New York: Basic Books.
- Parkes, C.M. (1965) Bereavement and mental illness. *British Journal of Medical Psychology*, 38, 1-16.
- Parkes, C.M. (1972) *Bereavement: Studies of Grief in Adult Life*. London: Pelican Books.
- Parkes, C.M. (1975) Determinants of outcome following bereavement. *Omega*, 6(4), 303-323.
- Parkes, C.M. (1988) Bereavement as a psychosocial transition: process of adaptation to change. *Journal Of Social Issues*, 44, 53-65.
- Parkes, C.M. (1995) Guidelines for conducting ethical bereavement research. *Death Studies*, 19, 171-181.
- Parrott, G. & Sabini, J. (1990) Mood and memory under natural conditions: Evidence for mood and incongruent recall. *Journal of Personality and Social Psychology*, 59, 321-336.
- Patton, M.Q. (1990) *Qualitative Evaluation and Research Methods*. 2<sup>nd</sup> ed. London: Sage.
- Pearlman, R.A. & Uhlmann, R.F. (1988) Quality of life in chronic diseases: perceptions of elderly patients. *Journal of Gerontology*, 43, M25-30.
- Pennebaker, J.W. (1983) Implicit psychophysiology: Effects of common beliefs and idiosyncratic physiological responses on symptom reporting. *Journal of Personality*, 51, 468-496.
- Pennebaker, J.W. (1984) Accuracy of symptom perception. In: A. Baum, S.E. Taylor & J.E. Singer (Eds.) *Handbook of psychology and health*. (Vol. IV), Hillsdale, NJ: Lawrence Erlbaum Associates. pp.189-217.
- Peruselli, C., Camporesi, E., Colombo, A.M., Cucci, M., Sironi, P.G., Bellodi, M., Cirillo, R., Love, E. & Mariano, R. (1992) Nursing care planning for terminally ill patients receiving home care. *Journal of Palliative Care*, 8(4), 4-7.
- Peteet, J., Tay, V., Cohen, G. & MacIntyre, J. (1986) Pain characteristics and treatment in an outpatient cancer population. *Cancer*, 57, 1259-1265.
- Pickle, L.W., Brown, L.M., Blot, W.J. (1983) Information available from surrogate respondents in case-control interview studies. *American Journal of Epidemiology*, 118, 99-108.
- Potkins, D., Bradley, S., Shrimanker, J., O'Brien, J., Swann, A. & Ballard, C. (2000) End of life treatment decisions in people with dementia: carers' views and the factors that influence them. *International Journal of Geriatric Psychiatry*, 15(11), 1005-1008.
- Potter, J. & Wetherell, M. (1987) *Discourse and Social Psychology: Beyond Attitudes and Behaviour*. London: Sage Publications.
- Presser, S. & Blair, J. (1994) Survey pretesting : Do different methods produce different results? In: P.V. Marsden (Ed.) *Sociological Methodology*, Vol.24. Washington, DC: American Statistical Association. pp.73-104.
- Punch, F. (1998) Mixed methods and evaluative criteria In: F. Punch (Ed.) *Introduction to Social Research*. London: Sage Publications. pp.239-263.



- Reading, A. (1982) A comparison of the McGill Pain Questionnaire in chronic and acute pain. *Pain*, 13, 185-198.
- Redinbaugh, E.M., Baum, A., DeMoss, C., Fello, M. & Arnold, R. (2002) Factors associated with the accuracy of family caregiver estimates of patient pain. *Journal of Pain and Symptom Management*, 23, 31-38.
- Rinck, G.C., van den Bos, G.A., Kleijnen, J., de Haes, H.J., Schade, E. & Veenhof, C.H. (1997) Methodologic issues in effectiveness research on palliative cancer care: a systematic review. *Journal of Clinical Oncology*, 15, 1697-1707.
- Robertson, J., Jay, J. & Welch, S. (1997) Can data collection during the grieving process be justifiable? *British Journal of Nursing*, 6, 759-765.
- Robinson, B.C. (1983) Validation of a Caregiver Strain Index. *Journal of Gerontology*, 38, 344-348.
- Rogers, J., Ridley, S., Chrispin, P., Scotton, H. & Llyod, D. (1997) Reliability of the next of kins' estimates of critically ill patients' quality of life. *Anaesthesia*, 52, 1137-1143.
- Rogot, E., & Reid, D.D. (1975) The validity of data from next-of-kin in studies of mortality among migrants. *International Journal of Epidemiology*, 4, 51-54.
- Rosenblatt, P.C. (1993) Grief: The social context of private feelings. In: M.S. Stroebe, W. Stroebe & R.O. Hansson (Eds.) *Handbook of bereavement: Theory, research, and intervention*. Cambridge, UK: Cambridge University Press. (pp.102-111)
- Ross, M. (1989) The relation of implicit theories to the construction of personal histories. *Psychological Bulletin*, 96, 341-357.
- Roth, A.J. & Breitbart, W. (1996) Psychiatric emergencies in terminally ill cancer patients. *Hematology - Oncology Clinics of North America*. 10(1), 235-259.
- Rothman, M.L., Hedrick, S.C., Bulcroft, K.A., Hickam, D.H. & Rubenstein, L.Z. (1991) The validity of proxy-generated scores as measures of patient health status., *Medical Care*, 4, 115-124.
- Royston, P.N. (1989) Using intensive interviews to evaluate questions. In: F.J. Fowler, Jr. (Ed.) *Health survey research methods*. Washington DC: Government Printing Office. pp.3-7.
- Rubenstein, L.Z., Schairer, C., Wieland, G.D. & Kane, R. (1984) Systematic biases in functional status assessment of elderly adults: effects of different data sources. *Journal of Gerontology*, 39, 686-691.
- Rubin, S.S. (1985) The resolution of bereavement: A clinical focus on the relationship to the deceased. *Psychotherapy*, 22(2), 231-235.
- Sales, E., Schulz, R. & Biegel, D. (1992) Predictors of strain in families of cancer patients: a review of the literature. *Journal of Psychosomatic Oncology*, 10, 1-26.
- Sandelowski, M. (1999) Time and Qualitative Research. *Research in Health & Nursing*, 22, 79-87.
- Sanders, C.M., Mauger, P.A. & Strong, P.A. (1985/1991) *A manual for the grief experience inventory*. Palo Alto, CA: Consulting Psychologists Press
- Scafeffer, N.C. (1991) Hardley ever or constantly? Group comparisons using vague quantifiers, *Public Opinion Quarterly*, 55, 395-423.
- Scherer, K.R. (1984) On the nature and function of emotion: a component process approach. In: K.R. Scherer & P. Ekman (Eds.) *Approaches to Emotion*. Hillsdale, NJ: Lawrence Erlbaum.

- Schipper, H. & Levitt, M. (1985) Measuring quality of life: risks and benefits. *Cancer Treatment Review*, 69, 1115-1123.
- Schrader, G., Davis, A., Stefanovic, S., & Christie, P. (1990) The recollection of affect. *Psychological Medicine*, 20, 105-109.
- Schulz, R., Visintainer, P. & Williamson, G.M. (1990) Psychiatric and physical morbidity effects of caregiving. *Journal of Gerontology: Psychological Sciences*, 45, 181-191.
- Schwarz, N. (1990) Feelings as information: Informational and motivational functions of affective states. In: E.T. Higgins & R.M. Sorrentino (Eds.) *Handbook of motivation and cognition: Foundations of social behavior*. New York: Guildford. Vol.2, pp.527-561.
- Schwarz, N. (1990) Assessing frequency Reports of Mundane Behaviors: Contributions of Cognitive Psychology to Questionnaire Construction. In: C. Hendrick & M.S. Clark (Eds.) *Research Methods in Personality and Social Psychology. (Review of Personality and Social Psychology, 11, 98-119)*. Newbury Park, CA: Sage.
- Schwarz, N. & Clore, G.L. (1988) How do I feel about it? Informative functions of affective states. In: K. Fieldler & J. Forgas (Eds.) *Affect, cognition and social behavior*. Toronto: Hofgreffe International. pp.44-62.
- Schwarz, N. & Clore, G.L. (1989) The informative function of affective states. In: F. Fiedler & J.P. Forgas (Eds.) *Affect, Cognition & Social Behavior*. Toronto: Hogrefe.
- Schwarz, N. & Sudman, S. (1996) *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco: Jossey-Bass.
- Seale, C. & Cartwright, A. (1994) *The Year before Death*. Aldershot: Avebury.
- Seamark, D.A., Lawrence, C. & Gilbert, J. (1996) Characteristics of referrals to an in-patient hospice and a survey of general practitioner perceptions of palliative care. *Journal of Royal Society of Medicine*, 89, 79-84.
- Seckler, A.B., Meier, D.E., Mulvihill, M. & Cammer Paris, B.E. (1991) Substituted Judgment: How Accurate Are Proxy Predictions? *Annals of Internal Medicine*, 115(2), 92-98.
- Segal, M.E. & Schall, R.R. (1994) Determining Functional/Health Status and Its Relation to Disability in Stroke Survivors. *Stroke*, 25(12), 2391-2397.
- Sensky, T., Dennehy, M., Gilbert, A., Begent, R., Newlands, E., Rustin, G. & Thompson, C. (1989) Physicians' perceptions of anxiety and depression among their outpatients: relationships with patients and their doctors' satisfaction with their interviews. *Journal of the Royal College of Physicians London*, 23, 33-38.
- Shakleton, C.H. (1984) The psychology of grief: a review. *Advances in Behaviour Research and Therapy*, 6, 153-205.
- Shuchter, S.R. & Zisook, S. (1993) The course of normal grief. In: M.S. Stroebe, W. Stroebe & R.O. Hansson (Eds.) *Handbook of bereavement: Theory, research, and intervention*. Cambridge, UK: Cambridge University Press. pp.23-43.
- Siegel, K., Raveis, V., Mor, V., Houts, P. (1991) The relationship of spousal caregiver burden to patient disease and treatment-related conditions. *Annals of Oncology*, 2, 511-516.
- Sigurdardottir, V., Brandberg, Y. & Sullivan, M. (1996) Criterion-based validation of the EORTC QLQ-C36 in advanced melanoma: the CIPS questionnaire and proxy raters. *Quality of Life Research*, 5, 375-386.

- Silver, R.L. & Wortman, C.B. (1980) Coping with undesirable life events. In: J. Garber & M.E.P. Seligman (Eds.) *Human helplessness: Theory and applications*. New York: Academic Press. pp.279-375.
- Silverfarb, Maurer, L.H. & Crouthamel, C.S. (1980) Psychosocial aspects of neoplastic disease: I. Functional status of breast cancer patients during different treatment regimes. *American Journal of Psychiatry*, 137, 450-455.
- Slevin, M.L., Plant, H., Lynch, D., Drinkwater, J. & Gregory, W.M. (1988) Who should measure quality of life, the doctor or the patient? *British Journal of Cancer*, 57, 109-112.
- Snaith, R.P. (1981) Rating scales. *British Journal of Psychiatry*, 138, 512-514.
- Sneeuw, K.C.A., Aaronson, N.K., de Haan, R.J. & Limburg, M. (1997c) Assessing quality of life after stroke: the value and limitations of proxy ratings. *Stroke*, 28, 1541-1549.
- Sneeuw, K.C., Aaronson, N.K., Osoba, D., Muller, M.J., Hsu, M.A., Yung, W.K., Brada, M. & Newlands, E.S. (1997a) The use of significant others as proxy raters of the quality of life of patients with brain cancer. *Medical Care*, 35, 490-506.
- Sneeuw, K.C., Aaronson, N.K., Sprangers, M.A., Detmar, S.B., Wever, L.D. & Schornagel, J.H. (1997b) Value of caregiver ratings in evaluating the quality of life of patients with cancer. *Journal of Clinical Oncology*, 15, 1206-1217.
- Sneeuw, K.C., Aaronson, N.K., Sprangers, M.A., Detmar, S.B., Wever, L.D. & Schornagel, J.H. (1998) Comparison of patient and proxy EORTC QLQ-C30 ratings in assessing the quality of life of cancer patients. *Journal of Clinical Epidemiology*, 51, 617-631.
- Sneeuw, K.C., Aaronson, N.K., Sprangers, M.A., Detmar, S.B., Wever, L.D. & Schornagel, J.H. (1999) Evaluation of the quality of life of cancer patients: assessments by patients, significant others, physicians and nurses. *British Journal of Cancer*, 81, 87-94.
- Sneeuw, K.C.A., Albertsen, P.C. & Aaronson, N.K. (2001) Comparison of patient and spouse assessments of health related quality of life in men with metastatic prostate cancer. *The Journal of Urology*, 165(2), 478-482.
- Somogyi-Zalud, E., Zhong, Z., Lynn, J. & Hamel, M.B. (2000) Elderly persons' last six months of life: findings from the hospitalized elderly longitudinal project. *Journal of the American Geriatrics Society*, 48, S131- S139.
- Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R. & Jacobs, G.A. (1983) *Manual for the State-Trait Anxiety Inventory for Adults (Revised Edition)*. Palo Alto, CA: Consulting Psychologists Press.
- Spiller, J.A. & Alexander, D.A. (1993) Domiciliary care: a comparison of the views of terminally ill patients and their family caregivers. *Palliative Medicine*, 7, 109-115.
- Sprangers, M.A. & Aaronson, N.K. (1992) The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease. *Journal of Clinical Epidemiology*, 45, 743-760.
- Srull, T.K. & Wyer, R.S. Jr. (1993) *The Mental Representation of Trait and Autobiographical Knowledge About the Self*. Hillsdale, NJ: Erlbaum.
- Stancliffe, R.J. (2000) Proxy respondents and quality of life. *Evaluation and Program Planning*, 23, 89-93.
- Stein, N.L., Folkman, S., Trabasso, T., & Christopher-Richards, A. (1997) Appraisal and goal processes as predictors of well-being in bereaved caregivers. *Journal of Personality and Social Psychology*, 72, 863- 871.

- Stein, N.L. & Levine, L.J. (1990) Making sense out of emotion: The representation and use of goal-structured knowledge. In: N.L. Stein, B. Leventhal & T. Trabasso (Eds.) *Psychological and biological approaches to emotion*. Hillsdale, NJ: Erlbaum. pp.45-73.
- Stein, N.L. & Trabasso, T. (1992) The Organisation of Emotional Experience: Creating Links among Emotion, Thinking, Language and Intentional Action. *Cognition and Emotion*, 6(3/4), 225-244.
- Sterkenburg, C.A., King, B. & Woodward, C.A. (1996) A reliability and validity study of the McMaster Quality of Life Scale (MQLS) for a palliative population. *Journal of Palliative Care*, 12, 18-25.
- Stez, K.M. (1987) Caregiving demands during advanced cancer: the spouse's needs. *Cancer Nursing*, 10, 260-268.
- Stone, P.J., Dunphy, D.C. & Smith, M.S. (1966) *The general inquirer: A computer approach to content analysis*. Cambridge, MA: MIT.
- Strauss, A. & Corbin, J. (1994) Grounded theory methodology: an overview. In: N.K. Denzin & Y.S. Lincoln (Eds.) *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage. pp.273-285.
- Stroebe, M.S., Hansson, R.O. & Stroebe, W. (1993) Themes in bereavement research. In: M.S. Stroebe, W. Stroebe & R.O. Hansson (Eds.) *Handbook of bereavement: Theory, research, and intervention*. Cambridge, UK: Cambridge University Press. pp.457-475.
- Stroebe, W. & Stroebe, M.S. (1993) Determinants of adjustment to bereavement in younger widows and widowers. In: M.S. Stroebe, W. Stroebe & R.O. Hansson (Eds.) *Handbook of bereavement: Theory, research, and intervention*. Cambridge, UK: Cambridge University Press. pp.208-226.
- Stroebe, W., Stroebe, M.S. & Domittner, G. (1985) The impact of recent bereavement on the mental and physical health of young widows and widowers. *Reports from the Psychological Institute of Tübingen University*.
- Sudman, S. & Bradburn, N.M. (1973) Effects of time and memory factors on response in surveys. *Journal of American Statistical Association*, 68, 805-815.
- Sudman, S. & Bradburn, N.M. (1974) *Response Effects in Surveys*. Chicago: Aldine.
- Sudman, S., Bradburn, N.M. & Schwarz, N. (1996) *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco: Jossey-Bass.
- Sulmasy, D.P., Terry, P.B., Weisman, C.S., Miller, D.B., Stallings, R.Y., Vettese, M.A. & Haller, K.B. (1998) The Accuracy of Substituted Judgment in Patients with Terminal Diagnoses. *Annals of Internal Medicine*, 128, 621-629.
- Taylor, S.E. (1991) Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, 110, 67-85.
- Teno, J.M., Clarridge, B., Casey, V., Edgman-Levitan, S. & Fowler, J. (2001) Validation of Toolkit After-Death Bereaved Family Member Interview. *Journal of Pain and Symptom Management*, 22(3), 752-758.
- Tepper, A., Connally, B., Haltmeier, P., Smith, E., & Sweeney, M.H. (1993) Knowledge of medical history information among proxy respondents for deceased subjects. *Journal of Clinical Epidemiology*, 46(11), 1243-1248.
- Teri, L. & Truax, P. (1994) Assessment of depression in Dementia Patients: Association of Caregiver Mood With Depression Ratings. *The Gerontologist*, 34(2), 231-234.
- Teri, L. & Wagner, A.W. (1991) Assessment of depression in Alzheimer's disease: Concordance between informants. *Psychology & Aging*, 6, 280-285.

- The Hospice Information Service (2000) *Directory of Hospice and Palliative Care Services in the United Kingdom and Republic of Ireland*. London: St Christopher's Hospice.
- Thompson, C.P. (1982) Memory for unique personal events: The roommate study. *Memory and Cognition*, 10, 324-332.
- Thompson, C.P., Skowronski, J.J., Larsen, S.F., & Betz, A. (1996) *Autobiographical memory: Remembering what and remembering when*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Tourangeau, R. (1984) Cognitive science and survey methods. In: T.B. Jabine, M.L. Straf, & R. Tourangeau, (Eds.) *Cognitive aspects of survey methodology: building a bridge between disciplines*. Washington DC: National Academy Press. pp.73-100
- Tourangeau, R. (1987) Attitude measurement: A cognitive perspective. In: H.J. Hippler, N.Schwarz, & S. Sudman, (Eds.) *Social information processing and survey methodology*. New York: Springer-Verlag. pp.149-162
- Tourangeau, R. & Rasinski, K.A. (1988) Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin*, 103, 299-314.
- Tsevat, J., Cook, E.F., Green, M.L., Matchar, D.B., Dawson, N.V., Broste, S.K., Wu, A.W., Phillips, R.S., Oye, R.K. & Goldman, L. (1995) Health values of the seriously ill. *Annals of Internal Medicine*, 122(7), 514-520.
- Tsevat, J., Dawson, M.D., Wu, A.W., Lynn, J., Soukup, J.R., Cook, E.F., Vidaillet, H. & Philips, R.S. (1998) Health values of hospitalized patients 80 years old or older. *JAMA*, 279, 371-375.
- Tversky, A. & Kahneman, D. (1973) Availability. A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.
- Tversky, A. & Kahneman, D. (1974) Judgment under uncertainty: heuristics and biases. *Science*, 125, 1124-1131.
- Uhlmann, R.F., Pearlman, R.A. & Cain, K.A. (1988) Physicians' and spouses' predictions of elderly patients' resuscitation preferences. *Journal of Gerontology*, 43, 115-121.
- Vachon, M.L.S., Rogers, J., Lyall, W.A., Lancee, W.A., Sheldon, A.R. & Freeman, S.J. (1982) Predictors and correlates of adaptation to conjugal bereavement. *American Journal of Psychiatry*, 139, 998-1002.
- Vachon, M. & Stylianos, S.K. (1988) The role of social support in bereavement. *Journal of Social Issues*, 44(3), 175-190.
- Vainio A. & Auvinen A. (1996) Prevalence of symptoms among patients with advanced cancer: an international collaborative study. *Journal of Pain and Symptom Management*. 12(1), 3-10.
- Ventafridda, V., Ripamonti, C., De Conno, F. & Tamburini, M. (1990) Symptom prevalence and control during cancer patients last days of life. *Journal of Palliative Care*, 6(3), 7-11.
- Vernon, M.D. (1970) *Perception through Experience*. London: Metbauen.
- Vigano, A., Dorgan, M., Buckingham, J., Bruera, E. & Suarez-Almazor, M.E. (2000) Survival prediction in cancer patients: a systematic review of the medical literature, *Palliative Medicine*, 14, 363-374.
- Volicer, L., Cantor, M.D., Derse, A.R., Edwards, D.M., Prudhomme, A.M., Gregory, D.C., Reagan, J.E., Tulskey, J.A. & Fox, E. (2002) Advanced care planning by proxy for residents of long-term care facilities who lack decision-making capacity. *Journal of the American Geriatrics Society*, 50(4), 761-767.

- Walker, W.R., Vogl, R.J. & Thompson, C.P. (1997) Autobiographical Memory: Unpleasantness Fades Faster Than Pleasantness Over Time. *Applied Cognitive Psychology*, 11, 399-413.
- Watkins, P.C., Matthews, A., Williams, D.A. & Fuller, R.D. (1992) Mood congruent memory in depression: emotional priming or elaboration? *Journal of Abnormal Psychology*, 101, 581-586.
- Watts, F.N. (1993) Problems of memory and concentration. In: C.G. Costello (ed.) *Symptoms of Depression*. New York: Wiley. pp.113-140.
- Weber, R.P. (1985) *Basic Content Analysis*. London: Sage Publications.
- Wegner, D.T. & Petty, R.E. (1996) The flexible correction model: The role of naive theories of bias in bias correction. In: M. Zanna (Ed.) *Advances in Experimental Psychology*, 29, 141-208.
- Weinberger, M., Samsa, G.P., Schmader, K. Greenberg, S.M., Carr, D.B. & Wildman, D.S. (1992) Comparing proxy and patients' perceptions of patients' functional status: Results from an outpatient geriatric clinic. *Journal of American Geriatrics Society*, 40(6), 585-588.
- Werner, P., Cohen-Mansfield, J., Watson, V. & Pasis, S. (1998) Pain in participants of adult day care centers: assessment by different raters. *Journal of Pain and Symptom Management*, 15, 8-17.
- WHO (1990) *Cancer Pain Relief and Palliative Care*. Geneva, Switzerland: World Health Organisation.
- Williams, J.M.G. & Dritschel, B.H. (1988) Emotional disturbance and the specificity of autobiographical memory. *Cognition and Memory*, 2, 221-234.
- Williams, J.M.G., Watts, F.N., Macleod, C. & Mathews, A. (1997) *Cognitive Psychology and Emotional Disorders* (2<sup>nd</sup> Ed.) Chichester: Wiley.
- Williams, R A., Hagerty, B.M., Cimprich, B., Therrien, B. Bay, E. & Oe, H. (2000) Changes in attention and short-term memory in depression. *Journal of Psychiatric Research*, 34(3), 227-238.
- Willis, G.B., Royston, P. & Bercini, D. (1991) The use of verbal report methods in the development and testing of survey questionnaires. *Applied Cognitive Psychology* 5, 251-267.
- Wilson, K.A., Dowling, A.J., Abdoell, M., Tannock, I.F. (2000) Perception of quality of life by patients, partners, and treating physicians. *Quality of Life Research*, 9, 1041-1052.
- Winkielman, P., Zajonc, R.B. & Schwarz, N. (1997) Subliminal affective priming resists attributional interventions. *Cognition and Emotion*, 11, 433-465.
- Winthrope, C. & Rabbitt, P.A. (1988) Working memory capacity, IQ, age and ability to recount autobiographical events. In: M.M.Gruenberg, P.E.Morris & R.N.Sykes (Eds.) *Practical Aspects of Memory: Current Research and Issues II*. Chichester: John Wiley. pp.175-179.
- Wortman, C.B. & Silver, R.C. (1992) Reconsidering assumptions about coping with loss: an overview of current research. In: L. Montada, S-H. Filipp & M.J. Lerner (Eds.) *Life Crises and Experiences of Loss in Adulthood*. Hillsdale, NJ: Erlbaum. pp.341-365.
- Wortman, C.B., Silver, R.C. & Kessler, R.C. (1999) The meaning of loss and adjustment to bereavement. In: M.S. Stroebe, W. Stroebe & R.O. Hansson (Eds.) *Handbook of Bereavement: Theory Research and Intervention*. Cambridge, UK: Cambridge University Press. pp.349-366.

- Wright, B.A. (1960) *Physical disability, a psychological approach*. New York: Harper & Row.
- Wright, B.A. (1983) *Physical disability, a psychosocial approach*. New York: Harper & Row.
- Wyer, R.S.Jr., Clore, G.L. & Isbell, L.M. (1999) Affect and Information Processing. In: M.P. Zanna (Ed.) *Advances in Experimental Social Psychology*. London: Academic Press. pp.1-77.
- Yeager, K.A., Miaskowski, C., Dibble, S.L. & Wallhagen, M. (1995) Differences in Pain Knowledge and Perception of the Pain Experience Between Outpatients With Cancer and Their Family Caregivers. *Oncology Nursing Forum*, 22, 1235-1241.
- Zanetti, O., Geroldi, C., Frisoni, G.B., Bianchetti, A. & Trabucchi, M. (1999) Contrasting results between caregiver's reports and direct assessment of activities of daily living in patients affected by mild and very mild dementia: The contribution of the caregiver's characteristics. *Journal of American Geriatric Society*, 47, 196-202.
- Zigmond, A.S. & Snaith. The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67, 361-370.
- Zimmerman, L. Story, K.T., Gaston-Story, F. & Rowles, J.R. (1996) Psychological variables and cancer pain. *Cancer Nursing*, 19(1), 44-53.
- Zisook, S., Schneider, D. & Shuchter, S.R. (1990) Anxiety and bereavement. *Psychiatric Medicine*, 8(2), 83-96.
- Zisook, S. & Schuchter, S.R. (1991) Depression through the first year after the death of a spouse. *American Journal of Psychiatry*, 148, 1346-1352.
- Zweibel, N.R. & Cassel, C.K. (1989) Treatment Choices at the End of Life: A Comparison of Decisions by Older Patients and Their Physician-Selected Proxies. *The Gerontologist*, 29(5), 615-621.

# List of Publications

## *Papers in scientific journals*

- McPherson, C.J. & Addington-Hall, J.M. (2003) Judging the quality of care at the end of life: can proxies provide reliable information? *Social Science & Medicine*, 56(1), 95-109.
- Addington-Hall, J. & McPherson, C. (2001) After-Death Interviews with Surrogates/bereaved family members: some issues of validity. *Journal of Pain and Symptom Management*, 22(3), 784-790.

## *Papers under review*

- McPherson, C.J. & Addington-Hall, J.M. The application of verbal protocol analysis to palliative care surveys: bereaved family members' evaluation of patients' pain, anxiety and depression. *Qualitative Health Research*.
- McPherson, C.J. & Addington-Hall, J.M. What do bereaved relatives understand by the pain, anxiety and depression descriptors used in palliative care surveys? *Palliative Medicine*.
- McPherson, C.J. & Addington-Hall, J.M. How do proxies' perceptions of patients' pain, anxiety and depression change during the bereavement period? *Journal of Palliative Care*.

## *Published abstracts*

- McPherson, C.J. & Addington-Hall, J.M. (2002) How do memories of a significant other's experience change during the bereavement period? 14<sup>th</sup> International Congress on Care of the Terminally Ill, Montreal, Canada. October.
- McPherson, C.J. & Addington-Hall, J.M. (2001) How do memories of a significant other's experience change during the bereavement period? (Preliminary findings) *British Psychological Society Centenary Annual Conference*, Glasgow, UK April.
- McPherson, C.J. & Addington-Hall, J.M. (2001) What do proxies' understand by pain descriptors used to assess terminally ill patient's pain retrospectively? Palliative Care Research Forum. Royal College of Physicians, London, UK. June.
- Addington-Hall, J.M. & Copp, C.<sup>4</sup> (1999) Retrospective studies of the last year of life: Methodological issues. *Book of Abstracts, 6th Congress of EAPC*, Geneva. p88.

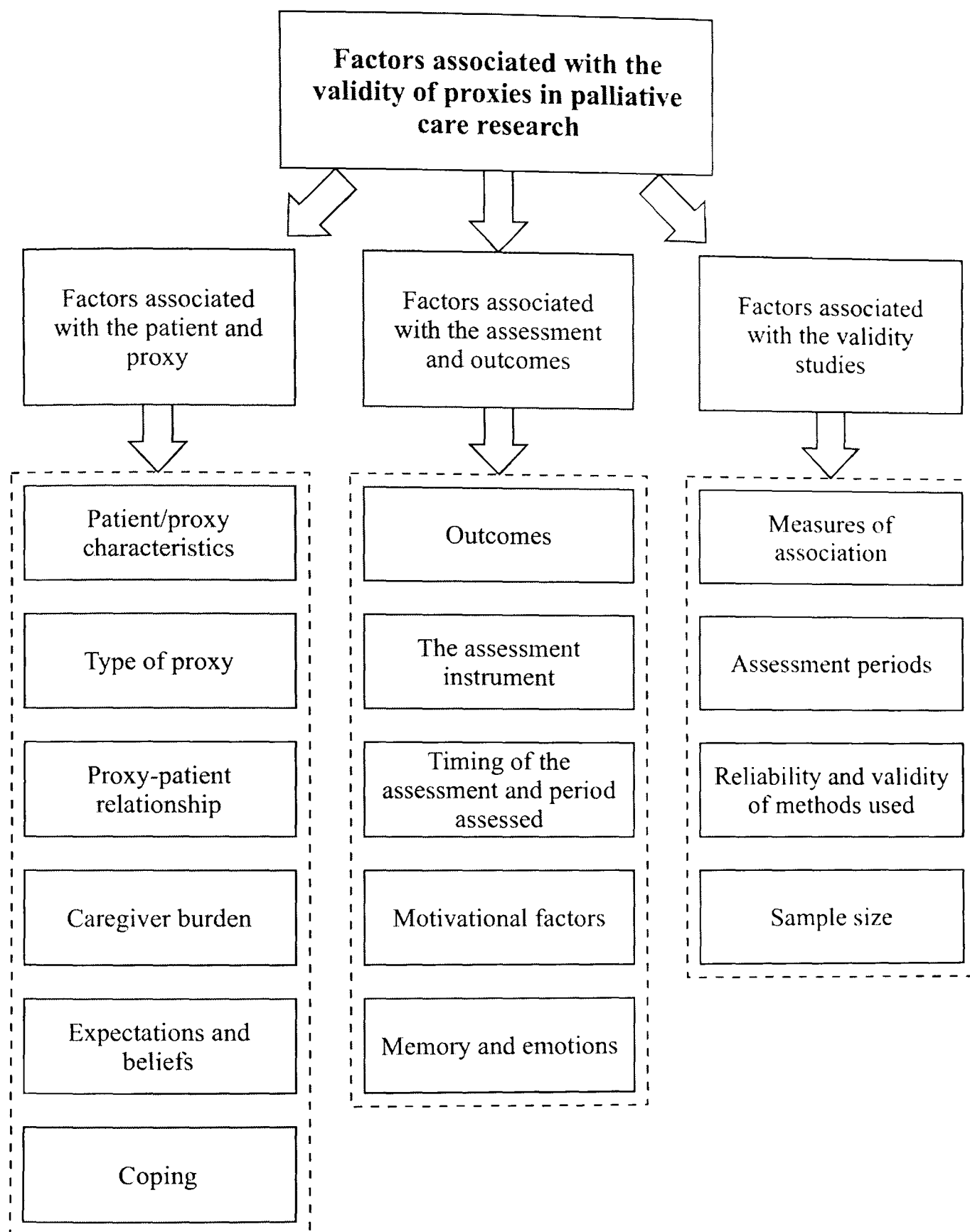
---

4. McPherson (previously Copp).



## **Appendix A**

Factors likely to affect the validity of proxies'  
retrospective reports



# Appendix B

## Ethics protocol

# Appendix B

## Ethics protocol

The following protocol was drawn up following negotiations with social workers at the hospice:

***Level One (All participants)***

All participants regardless of their psychological state will be given booklets and information on local bereavement and counselling services in their area. A full debriefing will be available to all participants, to allow them the opportunity to express their feelings about the interview.

***Level Two (Distressed participants)***

If concerned about an participant, for example their distress is unabated, ask if there is someone you can contact who can come and sit with them. Alternatively, ask them to contact their GP or give permission for you to contact their GP. Debriefing as above.

***Level Three (Very distressed participants)***

Participant showing signs of extreme distress (suicidal tendencies, hopelessness, and inability to cope). Ask for consent to contact their GP or a social worker. If they are unwilling to give consent contact the Social Work Department. Debriefing as above.

## **Appendix C**

Recruitment letter and participant consent for  
study one

*(Headed note paper)*

Dear \_\_\_\_\_ ,

I understand that you have recently experienced a bereavement and that your relative was in the care of ■■■■■. I realise that this is a very difficult time for you and I am sorry for any distress caused by this letter. The reason I am writing to you is to ask for your help with a study we are conducting. I will start by explaining the main reason for the research I am carrying out.

When patients are very ill they cannot always tell health professionals how they feel. So health professionals often rely on those closely involved with the patient, like yourself, to respond on behalf of the patient. It is really important to understand how you make decisions about the care they receive, as doctors and nurses use this information to plan the patient's care. Research studies so far show that what other people tell health professionals about the patients' feelings do not always match what the patient says themselves. There may be many reasons for this. One reason may be that the questions asked may mean different things to different people.

What we want to do is to ask some questions about three problems people who are very ill often experience; these are pain, anxiety and depression. The questions refer to the week and month before death. If you think that you are not the best person to answer these questions and there is someone else (a relative, friend or neighbour) who has had more contact with the patient in this period could you let me know. I appreciate these symptoms may be difficult to discuss, so I will do my best to deal with these issues as sensitively as possible. If necessary we can stop the interview at anytime.

The research is important because it can be used in the future to help plan and monitor the quality of care offered to terminally ill patients and their families and friends.

Your cooperation with the study is voluntary. All information you give will be strictly confidential and anonymous. It will in no way effect any services you may receive in future.

The study involves an interview with myself, Christine McPherson. I am a Registered Nurse studying in the Department of Palliative Care and Policy at King's College London. The interview will be carried out in your home or a quiet private place of your choice. The length of the interview will be decided by you. With your permission the interview will be tape-recorded so that an accurate record can be taken. Once the information has been obtained from these tapes, they will be destroyed.

I would be extremely grateful if you could help with the study as the information you provide will help us understand these symptoms from your viewpoint.

If you would like more information please do not hesitate to contact me on ■■■■■  
■■■■■ Monday, Wednesday or Thursday between 10.00 am - 7.00 pm.

If you do want to take part please fill in the form enclosed and return it in the stamped addressed envelope provided. Within 1-2 weeks I will contact you to arrange a convenient time to meet. If you do not want to take part you do not need to return the form.

Yours sincerely,

Christine McPherson

PhD Student/Registered Nurse

*(Headed note paper)*

Patient Identification Number:

**CONSENT FORM**

Title: Proxies retrospective perceptions of dying patients' pain, anxiety, and depression

*Please initial box*

1. I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions.

☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

☐

3. I agree to take part in the above study.

☐

\_\_\_\_\_  
*Name*

\_\_\_\_\_  
*Date*

\_\_\_\_\_  
*Signature*

Christine Jane McPherson

\_\_\_\_\_  
*Researcher*

\_\_\_\_\_  
*Date*

\_\_\_\_\_  
*Signature*



**Appendix D**

Demographic and patient/proxy characteristics

*The following questions are about your relationship to the deceased and information regarding their death.*

Can you tell me what your relationship is to the deceased?

- ☐ Spouse/partner
- ☐ Son/daughter
- ☐ Sibling
- ☐ In-law relationship
- ☐ Parent
- ☐ Neighbour/friend
- ☐ Other relationship (please state) \_\_\_\_\_

Can you tell me when he/she died? \_\_\_\_\_

Where did he/she die?

- ☐ Home
- ☐ Hospice
- ☐ Other (please state) \_\_\_\_\_

Where was he/she cared for in the week before he/she died?

- ☐ Home
- ☐ Hospice
- ☐ Other (please state)
- ☐ If more than one place (please state) \_\_\_\_\_

Can you tell me what he/she died of? \_\_\_\_\_

How old was he/she when he/she died? \_\_\_\_\_

Did you live with him/her? \_\_\_\_\_

How much contact did you have with him/her during their last week of life? \_\_\_\_\_

What was his/her previous occupation? \_\_\_\_\_

What was his/her ethnicity? \_\_\_\_\_

*The following are general questions about you.*

Gender: Male/ Female (circle)

Can you tell me how old you are? \_\_\_\_\_

What is your occupation/previous occupation? \_\_\_\_\_

What is your ethnicity? \_\_\_\_\_

## Appendix E

### Caregiver Strain Index

Caregiver Strain Index

Here is a list of things, which other people have found difficult when caring for someone. Could tick the box (yes/no) to indicate how you felt during the week before he/she died.

|    |  | Yes | No |
|----|--|-----|----|
| 1  | Sleep was disturbed  |     |    |
| 2  | It was inconvenient  |     |    |
| 3  | It was a physical strain   |     |    |
| 4  | It was constraining  |     |    |
| 5  | There were family adjustments  |     |    |
| 6  | There were changes in personal plans                                       |     |    |
| 7  | There were other demands on my time  |     |    |
| 8  | There were emotional adjustments   |     |    |
| 9  | Some behaviour was upsetting   |     |    |
| 10 | It was upsetting to find he/she had changed so much from their former self |     |    |
| 11 | There were work adjustments  |     |    |
| 12 | It was a financial strain  |     |    |
| 13 | I felt completely overwhelmed  |     |    |
|    | Total Score  |     |    |

Robinson, B.C. (1983) Validation of a Caregiver Strain Index. *Journal of Gerontology*, 38, 344-348.

## **Appendix F**

### **Proxy Semi-structured Cognitive Interview**

**Proxy Semi-structured Cognitive Interview**

***Introduction***

What I am going to do is to ask you questions about three problems people who are very ill sometimes experience, these are pain, anxiety and depression. There are no right or wrong answers, just say what comes to mind even if it seems out of place or trivial. Do not worry about trying to put your thoughts in any sort of order. We are really interested in what you have to say even if it may seem a bit muddled. Take your time, I will leave a pause after the questions for you to think. Just to remind you, if at any time you feel unable to respond or continue with the interview let me know and we can stop it at any time.

---

**Pain (Week)**

*Now I want to talk about any pain he/she might have experienced in the week before he/she died.*

1a. During the last week of life did \_\_\_\_\_ experience any pain?

- ☐ Yes
- ☐ No
- ☐ Don't know

*Prompt: was he/she uncomfortable on moving such as changing position or during activities, like dressing and washing (If no go to 1b; If don't know go to 1e)*

1b. In your own words, can you tell me what you were thinking of when you answered that question?  
(If no pain go to question 1d.)

1c. Were there any things that were more important than others in deciding whether he/she was in pain?

1d. Do you think a certain level of pain is expected given their condition? If so, can you explain?

1e. Can you tell me why you chose don't know?  
(Ask question 1d then go to question 2c.)

2a. How often did he/she experience pain in the last week of life?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

2b. When you answered question (2a.) how did you judge how often he/she was in pain?

2c. In your own words, what do you think to experience pain “often” means?

2d. Again in your own words, what do you think to experience pain “sometimes” means?

2e. What do you think to experience pain “rarely” means?  
(If no pain or don’t know go to 3c.)

3a. How intense was their pain during the last week of life?

- ☐ Severe pain
- ☐ Significant pain
- ☐ Moderate pain
- ☐ Mild pain
- ☐ None

3b. Can you tell me what were you thinking about when you answered that question?

In your own words, how would you describe:

3c. “Severe pain”?

3d. “Significant pain”?

3e. “Moderate pain”?

3f. “Mild pain”?

On a scale of pain severity 0 (painfree) to 10 (severe pain), where would you rate:

3g. “Significant pain”?

3h. “Moderate pain”?

3i. “Mild pain”?

(If no pain or don’t know go to 8a.)

4a. How distressing do you think he/she found the pain?

- ☐ Very distressing



- ☐ Fairly distressing
- ☐ Not very distressing
- ☐ Don't know

4c. What made you choose that option?

---

### **Pain relief in the last month**

5a. Did the doctor(s) give \_\_\_\_\_ any treatment to help relieve the pain?

- ☐ Yes
- ☐ No
- ☐ Not needed                      Go to 6e
- ☐ Don't know

5b. Who were you thinking of when I asked you that question?

6a. How effective was the pain control? Did it relieve the pain?

- ☐ Completely all the time
- ☐ Completely some of the time
- ☐ Partially
- ☐ Not at all
- ☐ Don't know

6b. What do you think "effective pain relief" means?

6c. Another response to choose from was "partially". What do you think this means in relation to pain relief?

6d. What were you thinking about when you answered that question?  
(Go to 7a.)

6e. Can you tell me why (you don't know or why he/she didn't receive help)?

---

7a. How distressing was this symptom to you?  
[On a scale of 0 to 10, 0 being 'no at all distressing' and 10 being 'extremely distressing']

---

## Pain (Month)

*Now I want to talk about any pain he/she might have experienced in the month before he/she died.*

8a. During the last month of life did \_\_\_\_\_ experience any pain?

☐ Yes

☐ No

☐ Don't know

*Prompt: was he/she uncomfortable on moving such as changing position or during activities, like dressing and washing (If no go to 8b; If don't know go to 8c.)*

8b. In your own words, can you tell me what you were thinking of when you answered that question?

(If no go to question 11a.)

8c. Can you tell me why you chose don't know or said no?

(Go to question 11a.)

9a. How often did he/she experience pain in the last month of life?

☐ Most of the time

☐ Often

☐ Sometimes

☐ Rarely

☐ Never

9b. When you answered that question (9a.) how did you judge how often he/she was in pain?

10a. How intense was his/her pain during the last month of life?

☐ Severe pain

☐ Significant pain

☐ Moderate pain

☐ Mild pain

☐ None

10b. What were you thinking of when you answered that question?

10c. On a scale of pain severity 0 (painless) to 10 (severe pain), where would you rate this?

---

**Anxiety (Week)**

*Changing the subject, I now want to talk about any anxiety or trouble with nerves he/she might have experienced in the week before he/she died.*

11a. During the last week of life did \_\_\_\_\_ experience any anxiety or trouble with nerves?

- ☐ Yes
- ☐ No
- ☐ Don't know
- Prompt: did he/she appear worried or concerned, express any fears (If no go to 11b; If don't know go to 11e.)*

11b. In your own words can you tell me what you were thinking about when you answered that question?  
(If no anxiety go to 11d.)

11c. Were there any signs that were more important than others in deciding whether he/she was anxious?

11d. Do you think a certain level of anxiety is expected given their condition? If so, can you tell me more about this?  
(If no anxiety go to 12c.)

11e. Can you tell me why you chose don't know?  
(Ask question 11d then go to question 12c.)

12a. How often did he/she experience anxiety?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely

☐ Never

12b. How did you decide how often he/she was anxious?

12c. In your own words, what do you think to experience anxiety “often” means?

12d. Again in your own words, what do you think to experience anxiety “sometimes” means?

12e. What do you think to experience anxiety “rarely” means?  
(If no anxiety or don’t know go to 13c.)

13a. How intense was their level of anxiety during the last week of life?

- ☐ Severe anxiety
- ☐ Significant anxiety
- ☐ Moderate anxiety
- ☐ Mild anxiety
- ☐ None

13b. How did you decide how intense their anxiety was?

In your own words what comes to mind when you think of:

- 13c. “Severe anxiety”?
- 13d. “Significant anxiety”?
- 13e. “Moderate anxiety”?
- 13f. “Mild anxiety”?

On a scale of anxiety severity 0 (no anxiety) to 10 (severe anxiety), where would you rate:

- 13g. “Significant anxiety”?
- 13h. “Moderate anxiety”?
- 13i. “Mild anxiety”?

(If no anxiety or don’t know go to 16a.)

---

# Help to relieve anxiety

14a. Did he/she get any help to cope with these problems?

- ☐ Yes
- ☐ No
- ☐ No help needed                      Go to question 14f.
- ☐ Don't know

14b. What help did he/she receive?

- ☐ Counselling
- ☐ Medication
- ☐ Other (please state)

14c. Can you tell me some more about the help he/she received?

14d. Did the help relieve the anxiety?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

14e. Can you tell me why you chose that option?  
(Go to 15a.)

14f. Can you tell me why (you don't know or why he/she didn't receive help)?

---

15a. How distressing was this symptom to you?

[On a scale of 1 to 10, 0 being 'not at all distressing' and 10 being 'extremely distressing']

---

## Anxiety (Month)

*Changing the subject, I now want to talk about any anxiety or trouble with nerves he/she might have experienced in the month before he/she died.*

16a. During the last month of life did \_\_\_\_\_ experience any anxiety or trouble with nerves?

☐ Yes

☐ No

☐ Don't know

*Prompt: did he/she appear worried or concerned, express any fears (If no go to 16b; If don't know go to 16c.)*

16b. In your own words can you tell me what you were thinking about when you answered that question?  
(If no anxiety go to question 19a.)

16c. Can you tell me why you chose don't know?  
(Go to question 19a.)

17a. How often did he/she experience anxiety in the month before death?

☐ Most of the time

☐ Often

☐ Sometimes

☐ Rarely

☐ Never

17b. When you answered that question (17a.) how did you judge how often he/she was anxious?

18a. How intense was their level of anxiety during the last month of life?

☐ Severe anxiety

☐ Significant anxiety

☐ Moderate anxiety

☐ Mild anxiety

☐ None

18b. How did you decide how intense his/her anxiety was?

18c. On a scale of anxiety severity 0 (no anxiety) to 10 (severe anxiety), where would you rate this?

---

**Depression (Week)**

*Changing the subject, I now want to talk about any depression he/she might have experienced in the week before he/she died.*

19a. During the last week of life did \_\_\_\_\_ experience any depression?

- ☐ Yes
- ☐ No *Prompt: did he/she appear down, sad, withdrawn or tearful (If no go to 19b.; If don't know go to 19e.)*
- ☐ Don't know

19b. In your own words can you tell me what you were thinking of when you answered that question?  
(If no depression go to question 19d.)

19c. Were there any signs that were more important than others in deciding whether he/she was depressed?

19d. Do you think a certain level of depression is expected given their condition? If so, can you tell me more about this?  
(If no depression go to question 20c.)

19e. Can you tell me why you chose don't know?  
(Ask question 19d. then go to question 20c.)

20a. How often did he/she experience depression in the last week of life?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

20b. When you answered that question (20a.) how did you judge how often he/she was depressed?

20c. In your own words, what do you think to experience depression “often” means?

20d. In your own words, what do you think to experience depression “sometimes” means?

20e. In your own words, what do you think to experience depression “rarely” means?  
(If no depression or don’t know go to question 21c.)

21a. How intense was his/her level of depression during the last week of life?

- ☐ Severe depression
- ☐ Significant depression
- ☐ Moderate depression
- ☐ Mild depression
- ☐ None

21b. What were you thinking of when you answered that question?

In your own words what comes to mind when you think of :

21c. “Severe depression”?

21d. “Significant depression”?

21e. “Moderate depression”?

21f. “Mild depression”?

On a scale of depression severity 0 (no depression) to 10 (severe depression), where would you rate:

21g. “Significant depression”

21h. “Moderate depression”

21i. “Mild depression”

(If no depression or don’t know go to question 24a.)



**Help to relieve depression**

22a. Did he/she get any help to cope with these problems?

- ☐ Yes
- ☐ No
- ☐ No help needed                      Go to question 22f.
- ☐ Don't know

22b. What help did he/she receive?

- ☐ Counselling
- ☐ Medication
- ☐ Other (please state)

22c. Can you tell me some more about the help he/she received?

22d. Did the help relieve their depression?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

22e. Can you tell me why you chose that option?  
(Go to 23a.)

22f. Can you tell me why (you don't know or why he/she didn't receive help)?

23a. How distressing was this symptom to you?  
[On a scale of 1 to 10, 0 being 'not at all distressing' and 10 being 'extremely distressing']

---

**Depression (Month)**

*Changing the subject, I now want to talk any depression he/she might have experienced in the month before he/she died.*

24a. During the last month of life did \_\_\_\_\_ experience any depression?

- ☐ Yes
- ☐ No
- ☐ Don't know

*Prompt: did he/she appear down, sad, withdrawn or  
tearful (If no go to 24b; If don't know go to 24c.)*

24b. In your own words can you tell me what you were thinking of when you answered that question?  
(If no depression go to question 27a.)

24c. Can you tell me why you chose don't know?  
(Go to question 27a.)

25a. How often did he/she experience depression during the last month of life?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

25b. When you answered that question (25a.) how did you judge how often he/she was depressed?

26a. How intense was his/her level of depression during the last month of life?

- ☐ Severe depression
- ☐ Significant depression
- ☐ Moderate depression
- ☐ Mild depression
- ☐ None

26b. How did you decide how intense his/her level of depression was?

26c. On a scale of depression severity 0 (no depression) to 10 (severe depression), where would you rate this?

---

27a. If I asked you in the week before death do you think your answers would be different. If so, how?

28a. If I asked you in the month after death do you think your answers would be different. If so, how?

---

## **Appendix G**

Differences between questions in the proxy  
semi-structured cognitive interview and  
questions from the retrospective palliative care  
questionnaires

*Differences between questions in the PSCI and questions in the retrospective palliative care questionnaires investigated.*

1. VOICES and the RSCD questionnaire assess symptoms over a period of twelve months, whereas the PSCI assesses questions relating to the last month and week of life. For example, VOICES asks

*During the last twelve months of his/her life, while he/she was at home, or while in the nursing or residential home, did he/she suffer any pain?*

Yes    ☐1                  No    ☐2                  Don't know    ☐3

2. One of the points raised in the introduction was the length of the assessment period. Although assessment of the last year may be useful in auditing services and information, symptoms can vary enormously during such an extended period of time. Consequently, it was decided that the PSCI's period of assessment would be in the later stages of patient's illness, when more reliance may be placed on others to assess patient's symptoms.
3. VOICES collapses anxiety and depression into one question, "*During his last year, did \_\_\_\_\_ suffer from anxiety and/or nerves or get depressed*"? Although the question is useful for assessing the psychosocial care of the patient, one of the roles of audit is to identify specific areas where improvements can be made. By differentiating between the anxiety and depression, treatments and care can be better targeted. There is also the danger of a proxy responding "*no*" because the patient did not have anxiety or because they did not have depression even though they did experience one of them. By separating out the two conditions it should also be possible to assess whether proxies have difficulty differentiating. In the end, it might turn out that the two conditions are in fact better combined.
4. VOICES does not provide severity scales for pain, anxiety or depression, yet it does for other symptoms.

|                          | A lot                      | Some                       | A little                   | Not at all                 | Don't know                 | Not treated                |
|--------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| Vomiting or feeling sick | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 | <input type="checkbox"/> 3 | <input type="checkbox"/> 4 | <input type="checkbox"/> 5 | <input type="checkbox"/> 6 |

5. In the Regional Study of the Care of the Dying upon which VOICES is based, pain severity was assessed by the amount of distress it caused patients. As symptom distress focuses on the affective elements of a symptom that are very subjective aspects. It was thought that this might be difficult to assess given the potential biases highlighted in the introduction. Therefore, an alternative was a symptom severity scale similar to the one adopted in another retrospective survey conducted by the developers of VOICES.
6. Frequency is assessed for anxiety/depression (as shown below) in VOICES, but not for pain. The RSCD assesses the length of time the patient had pain, for example, 3 months, 6 months. This gives an indication of the presence or absence of pain over a given period but does not assess the frequency when pain is present. Therefore, the same frequency response scale used to assess anxiety/depression, was used to assess pain frequency in this study.

During his last year, did \_\_\_\_\_ suffer from anxiety and/or nerves or get depressed?

Please tick one box

- |                       |                            |
|-----------------------|----------------------------|
| Yes, most of the time | <input type="checkbox"/> 1 |
| Yes, often            | <input type="checkbox"/> 2 |
| Yes, sometimes        | <input type="checkbox"/> 3 |
| Rarely                | <input type="checkbox"/> 4 |
| Never                 | <input type="checkbox"/> 5 |

## Prompts used in the Proxy Semi-structured Cognitive Interview

## Prompts used in the Proxy Semi-structured Cognitive Interview

*Question ambiguity*

Prompts were kept to a minimum in the interest of maximising participant’s interpretation of questions. If they were unsure of the question, it would be repeated. Care was taken to avoid “cueing” responses that reflected those of the interviewer instead of those of the participants. Participants often looked to the interviewer for confirmation that their replies were adequate or commented on their ability to provide the right information. To avoid reinforcing particular responses, the interviewer stressed that there were no right or wrong answers and that it was the participants’ views that were of interest.

*Timing of events*

In some instances, the participant appeared to talk about periods outside the periods being assessed. When they did so, they were asked again about the timing of the events. For example, when asked about the last week or month of life, participants often made reference to events in the last few months prior to the patient’s death. In order to clarify the period, the interviewer would ask “*Was that in the last week or month*”? When references were made to places of care, such as “*In the hospital...*,” the interviewer inquired about the time period this covered. This was an indirect method of ascertaining the timing of events.

*Symptoms*

In a number of instances when asked about the severity and frequency of their relative’s symptoms, the participant was unsure about whether the symptom was rated in one or another of the response options. This often occurred when there was variability in the symptom’s severity or frequency over time. When this happened participants were asked: “*If they were sent the questionnaire to complete, how they would respond?*” For example, would they choose both or one or none of the response options and why.

Some prompts were included in the PSCI, for example,

16a. During the last month of life did \_\_\_\_\_ experience any anxiety or trouble with nerves?

- ☐ Yes
- ☐ No
- ☐ Don’t know
- Prompt: did he/she appear worried or concerned, express any fears (If no go to 16b; If don’t know go to 16c.)*

These prompts were used when participants answered “don’t know” or “no” when asked about the presence or absence of a symptom. The reason for this, was to examine whether participants were answering “*don’t know*” or “*no*” because they didn’t consider that the patient had the symptom as it wasn’t severe enough, or because they were unsure about the symptom



in question. Even after using these prompts, none of the participants changed their responses to the question.

## **Appendix I**

Modifications to the Proxy Semi-structured  
Cognitive Interview following the pilot study

The pilot study resulted in few modifications to the interview. The changes were to the PSCI and VOICES.

1. The most significant change was to the use of the word “*suffer*” in VOICES. This is used to assess patients’ symptoms, for example, “*While he was in that hospice, did he suffer any pain related to his final illness*”? In the PSCI the question was “*During the last week of life did \_\_\_\_\_ suffer any pain*”? Participants in the pilot study found that the wording had implications regarding the responsibility of the patient’s symptom. As those interviewed were the main caregiver this had emotional undertones of blame for not controlling the symptom. One might also argue that suffer might not be an appropriate term to use in cases where a symptom is mild and the patient is not seen to suffer. Therefore, the word suffer was replaced with experience as this was considered a neutral term.
2. Originally PSCI included items assessing the distress caused by anxiety and depression. As these symptoms encompass distress within the symptomology it was thought that these questions were redundant. As one participant pointed out,  
*“Anxiety is distress isn’t it? ... so how do I know whether it caused her distress when it’s the same thing?”*
3. During the follow up study one participant commented on how difficult it was to discuss these issues now in comparison to the earlier interview. This initiated the inclusion of the question “*How does being interviewed now compare to the previous interview*”? at the end of the follow up interview.

## **Appendix J**

Recruitment letter and participant consent for  
the follow up study

*(Headed note paper)*

Dear \_\_\_\_\_ ,

My name is Christine McPherson, I interviewed you several weeks ago as part of a research study I was conducting. The interview asked for your views on the care of your deceased relative. In particular, any pain, anxiety and depression he/she may have experienced during their illness. I would like this opportunity to thank you again for your assistance with this research. I realised that this was a very difficult matter to talk about.

The reason I am writing to you is to ask for your help with a follow up study. The study will involve being interviewed again by myself. I am a Registered Nurse studying in the Department of Palliative Care and Policy at King's College London. The interview will be carried out in your home or a quiet private place of your choice. The interview will be much shorter than the first interview. It is estimated to take approximately half an hour to conduct. The questions will be about your relative's care and any pain, anxiety and depression he/she might have experienced.

This follow up study is important in understanding these symptoms from your viewpoint. It can be used in the future to help plan and monitor the quality of care offered to terminally ill patients and their families and friends.

Your cooperation with the study is voluntary. All information you give will be strictly confidential and anonymous. It will in no way effect any services you may receive in future.

If you would like more information please do not hesitate to contact me on ■■■■■■ ■■■■■■. If you do want to take part please fill in the form enclosed and return it in the stamped addressed envelope provided. Within 1-2 weeks I will contact you to arrange a convenient time to meet. If you do not want to take part you do not need to return the form.

Yours sincerely,

Christine McPherson

Ph.D. Student/Registered Nurse

(Headed note paper)

Patient Identification Number:

**CONSENT FORM**

Title: Proxies retrospective perceptions of dying patients' pain, anxiety and depression

*Please initial box*

1. I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions.

☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

☐
3. I agree to take part in the above study.

☐

|                              |             |                  |
|------------------------------|-------------|------------------|
| _____                        | _____       | _____            |
| <i>Name</i>                  | <i>Date</i> | <i>Signature</i> |
| <br>Christine Jane McPherson |             |                  |
| _____                        | _____       | _____            |
| <i>Researcher</i>            | <i>Date</i> | <i>Signature</i> |

## **Appendix K**

Shortened version of the Proxy Semi-structured  
Cognitive Interview used in the follow up study

**Proxy Semi-structured Cognitive Interview (Shortened Version).**

***Introduction***

What I am going to do is to ask you questions about three problems people who are very ill sometimes experience, these are pain, anxiety and depression. There are no right or wrong answers, just say what comes to mind even if it seems out of place or trivial. Do not worry about trying to put your thoughts in any sort of order. We are really interested in what you have to say even if it may seem a bit muddled. Take your time, I will leave a pause after the questions for you to think. Just to remind you, if at any time you feel unable to respond or continue with the interview let me know and we can stop it at any time.

---

**Pain (Week)**

*Now I want to talk about any pain he/she might have experienced in the week before he/she died*

1a. During the last week of life did \_\_\_\_\_ experience any pain?

- ☐ Yes
- ☐ No
- ☐ Don't know

*Prompt: was he/she uncomfortable on moving such as changing position or during activities, like dressing and washing (If no go to 1b; If don't know go to 1e)*

1b. In your own words, can you tell me what you were thinking of when you answered that question?  
(If no pain go to question 8a.)

1e. Can you tell me why you chose don't know?  
(Ask question 1d then go to question 8a.)

2a. How often did he/she experience pain in the last week of life?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

2b. When you answered question (2a.) how did you judge how often he/she was in pain?



3a. How intense was their pain during the last week of life?

- ☐ Severe pain
- ☐ Significant pain
- ☐ Moderate pain
- ☐ Mild pain
- ☐ None

3b. Can you tell me what were you thinking about when you answered that question?

4a. How distressing do you think he/she found the pain?

- ☐ Very distressing
- ☐ Fairly distressing
- ☐ Not very distressing
- ☐ Don't know

4c. What made you choose that option?

---

### **Pain relief in the last month**

5a. Did the doctor(s) give \_\_\_\_\_ any treatment to help relieve the pain?

- ☐ Yes
- ☐ No
- ☐ Not needed                      Go to 6e.
- ☐ Don't know

5b. Who were you thinking of when I asked you that question?

6a. How effective was the pain control? Did it relieve the pain?

- ☐ Completely all the time
- ☐ Completely some of the time
- ☐ Partially
- ☐ Not at all

☐ Don't know

6d. What were you thinking about when you answered that question?  
(Go to 7a.)

6e. Can you tell me why (you don't know or why he/she didn't receive help)?

---

7a. How distressing was this symptom to you?  
[On a scale of 0 to 10, 0 being 'no at all distressing' and 10 being 'extremely distressing']

---

---

**Pain (Month)**

*Now I want to talk about any pain he/she might have experienced in the month before he/she died.*

8a. During the last month of life did \_\_\_\_\_ experience any pain?

- ☐ Yes
  - ☐ No
  - ☐ Don't know
- Prompt: was he/she uncomfortable on moving such as changing position or during activities, like dressing and washing (If no go to 8b.; If don't know go to 8c.)*

8b. In your own words, can you tell me what you were thinking of when you answered that question?  
(If no go to question 11a.)

8c. Can you tell me why you chose don't know or said no?  
(Go to question 11a.)

9a. How often did he/she experience pain in the last month of life?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

9b. When you answered that question (9a.) how did you judge how often he/she was in pain?

10a. How intense was his/her pain during the last month of life?

- ☐ Severe pain
- ☐ Significant pain
- ☐ Moderate pain
- ☐ Mild pain
- ☐ None

10b. What were you thinking of when you answered that question?

10c. On a scale of pain severity 0 (painless) to 10 (severe pain), where would you rate this?

---

### Anxiety (Week)

*Changing the subject, I now want to talk about any anxiety or trouble with nerves he/she might have experienced in the week before he/she died.*

11a. During the last week of life did \_\_\_\_\_ experience any anxiety or trouble with nerves?

- ☐ Yes
- ☐ No
- ☐ Don't know

*Prompt: did he/she appear worried or concerned, express any fears (If no go to 11b; If don't know go to 11e.)*

11b. In your own words can you tell me what you were thinking about when you answered that question?  
(If no anxiety go to 17a.)

11e. Can you tell me why you chose don't know?  
(Ask question 11d. then go to question 17a.)

12a. How often did he/she experience anxiety?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

12b. How did you decide how often he/she was anxious?

12c. In your own words, what do you think to experience anxiety “often” means?

13a. How intense was their level of anxiety during the last week of life?

- ☐ Severe anxiety
- ☐ Significant anxiety
- ☐ Moderate anxiety
- ☐ Mild anxiety
- ☐ None

13b. How did you decide how intense their anxiety was?

15a. How distressing was this symptom to you?

[On a scale of 1 to 10, 0 being ‘not at all distressing’ and 10 being ‘extremely distressing’]

---

**Anxiety (Month)**

*Changing the subject, I now want to talk about any anxiety or trouble with nerves he/she might have experienced in the month before he/she died.*

16a. During the last month of life did \_\_\_\_\_ experience any anxiety or trouble with nerves?

- ☐ Yes
- ☐ No
- ☐ Don’t know

*Prompt: did he/she appear worried or concerned, express any fears (If no go to 16b.; If don’t know go to 16c.)*

16b. In your own words can you tell me what you were thinking about when you answered this question?  
(If no go to question 19a.)

16c. Can you tell me why you chose don’t know?  
(Go to question 19a.)

17a. How often did he/she experience anxiety in the month before death?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

17b. When you answered that question (17a.) how did you judge how often he/she was anxious?

18a. How intense was their level of anxiety during the last month of life?

- ☐ Severe anxiety
- ☐ Significant anxiety
- ☐ Moderate anxiety
- ☐ Mild anxiety
- ☐ None

18b. How did you decide how intense his/her anxiety was?

18c. On a scale of anxiety severity 0 (no anxiety) to 10 (severe anxiety), where would you rate this?

---

## Depression (Week)

*Changing the subject, I now want to talk about any depression he/she might have experienced in the week before he/she died.*

19a. During the last week of life did \_\_\_\_\_ experience any depression?

- ☐ Yes
- ☐ No
- ☐ Don't know

*Prompt: did he/she appear down, sad, withdrawn or tearful (If no go to 19b.; If don't know go to 19e.)*

19b. In your own words can you tell me what you were thinking of when you answered that question?  
(If no depression go to question 27a.)

19e. Can you tell me why you chose don't know?  
(Ask question 19d then go to question 27a.)

20a. How often did he/she experience depression in the last week of life?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

20b. When you answered that question (20a.) how did you judge how often he/she was depressed?

21a. How intense was his/her level of depression during the last week of life?

- ☐ Severe depression
- ☐ Significant depression
- ☐ Moderate depression
- ☐ Mild depression
- ☐ None

21b. What were you thinking of when you answered that question?

23a. How distressing was this symptom to you?  
[On a scale of 1 to 10, 0 being '*not at all distressing*' and 10 being '*extremely distressing*']

---

## Depression (Month)

*Changing the subject, I now want to talk any depression he/she might have experienced in the month before he/she died.*

24a. During the last month of life did \_\_\_\_\_ experience any depression?

- ☐ Yes
- ☐ No
- ☐ Don't know

*Prompt: did he/she appear down, sad, withdrawn or  
tearful (If no go to 24b.; If don't know go to 24c.)*

24b. In your own words can you tell me what you were thinking of when you answered that question?

(If no depression go to question 29a.)

24c. Can you tell me why you chose don't know?

(Go to question 29a.)

25a. How often did he/she experience depression during the last month of life?

- ☐ Most of the time
- ☐ Often
- ☐ Sometimes
- ☐ Rarely
- ☐ Never

25b. When you answered that question (25a.) how did you judge how often he/she was depressed?

26a. How intense was his/her level of depression during the last month of life?

- ☐ Severe depression
- ☐ Significant depression
- ☐ Moderate depression
- ☐ Mild depression
- ☐ None

26b. How did you decide how intense his/her level of depression was?

26c. On a scale of depression severity 0 (no depression) to 10 (severe depression), where would you rate this?

---

---

29a. How does being interviewed now compare to the previous interview?

---

---



**Appendix L**

Questions relating to symptom severity and  
frequency descriptors

## Pain

### *Frequency*

- 2c. In your own words, what do you think to experience pain “often” means?
- 2d. Again in your own words, what do you think to experience pain “sometimes” means?
- 2e. What do you think to experience pain “rarely” means?

### *Severity*

- 3c. How would you describe “severe pain”?
- 3d. How would you describe “significant pain”?
- 3e. How would you describe “moderate pain”?
- 3f. How would you describe “mild pain”?

### *Numerical rating*

On a scale of pain severity 0 (painfree) to 10 (severe pain), where would you rate:

- 3g. “Significant pain”
- 3h. “Moderate pain”
- 3i. “Mild pain”

## Anxiety

### *Frequency*

- 12c. In your own words, what do you think to experience anxiety “often” means?
- 12d. Again in your own words, what do you think to experience anxiety “sometimes” means?
- 12e. What do you think to experience anxiety “rarely” means?

### *Severity*

In your own words what comes to mind when you think of:

- 13c. “Severe anxiety”?
- 13d. “Significant anxiety”?
- 13e. “Moderate anxiety”?
- 13f. “Mild anxiety”?

### *Numerical rating*

On a scale of anxiety severity 0 (no anxiety) to 10 (severe anxiety), where would you rate:

13g. “Significant anxiety”

13h. “Moderate anxiety”

13i. “Mild anxiety”

## Depression

### *Frequency*

20c. In your own words, what do you think to experience depression “often” means?

20d. In your own words, what do you think to experience depression “sometimes” means?

20e. In your own words, what do you think to experience depression “rarely” means?

### *Severity*

In your own words what comes to mind when you think of:

21c. “Severe depression”?

21d. “Significant depression”?

21e. “Moderate depression”?

21f. “Mild depression”?

### *Numerical rating*

On a scale of depression severity 0 (no depression) to 10 (severe depression), where would you rate:

21g. “Significant depression”

21h. “Moderate depression”

21i. “Mild depression”

## **Appendix M**

Differences between individuals who took part  
in the study and those who did not

|                      | Took part | Did not take part | Test statistic                               | Significance     | <i>p</i> |
|----------------------|-----------|-------------------|--|------------------|----------|
| <i>Patient age</i>   |           |                   |  |                  |          |
| <i>N</i>             | 30        | 56                | <i>t</i> -test independent sample (2 tailed) | <i>t</i> = -0.97 | 0.92     |
| Range                | 25-90     | 32-94             |  |                  |          |
| Mean                 | 71        | 71                |  |                  |          |
| STD                  | 13.41     | 13.16             |  |                  |          |
| <i>Relationship</i>  |           |                   |  |                  |          |
| Spouse               | 20        | 33                | $\chi^2$ , df = 1                            | $\chi^2$ = .49   | 0.48     |
| Not spouse           | 10        | 26                |  |                  |          |
| <i>Patient Death</i> |           |                   |  |                  |          |
| Home                 | 11        | 13                | $\chi^2$ , df = 2                            | $\chi^2$ = 3.85  | 0.15     |
| Hospice              | 14        | 23                |  |                  |          |
| Hospital/not home    | 5         | 20                |  |                  |          |
| <i>Patient sex</i>   |           |                   |  |                  |          |
| Male                 | 15        | 32                | $\chi^2$ , df = 1                            | $\chi^2$ = .40   | 0.53     |
| Female               | 15        | 24                |  |                  |          |
| <i>Proxy sex</i>     |           |                   |  |                  |          |
| Male                 | 11        | 17                | $\chi^2$ , df = 1                            | $\chi^2$ = .35   | 0.55     |
| Female               | 19        | 39                |  |                  |          |

Two tailed,  $\alpha$  = .05.

Note: The relationship variable was collapsed to enable analysis, as some groupings contained frequencies of only 1 or 2. There was a greater variety of relatives in the non-participant group including several not found in those who took part, e.g. siblings, parent, aunt, and friends. The complete breakdown is presented below:

| Took part    | <i>N</i> | Did not take part | <i>N</i> |
|--------------|----------|-------------------|----------|
| Partner      | 18       | Partner           | 33       |
| Son/daughter | 10       | Son/daughter      | 5        |
| Ex-partner   | 2        | Sibling           | 14       |
|              |          | Mother            | 1        |
|              |          | Aunt              | 1        |
|              |          | Friend            | 2        |

**Appendix N**  
Coding frequencies of the Proxy Semi-structured Cognitive Interview

Pain coding

| Code description  | Week |     |     |     | Month |     |      |
|---|------|-----|-----|-----|-------|-----|------|
|   | 1b.  | 1c. | 2b. | 3b. | 8b.   | 9b. | 10b. |
|   |      |     |     |     |       |     |      |
| <b>Response difficulties</b>  |      |     |     |     |       |     |      |
| <i>Indicates difficulties proxies might experience when answering a question</i>  |      |     |     |     |       |     |      |
| No verbalisation of strategy  | 1    |     |     |     |       |     |      |
| No apparent strategy (e.g. guess)   |      |     | 1   |     |       |     |      |
| Don't know  |      | 1   |     | 2   |       |     |      |
| Incoherent response   |      |     |     |     |       |     |      |
| Explicitly states that the question is difficult to answer no reason given why  |      |     | 2   | 2   |       |     | 2    |
| States that the question is difficult to answer because of patient's level of consciousness (e.g. due to medication or confusion)   | 3    | 1   | 1   | 4   |       |     |      |
| States that the question is difficult to answer because of lack of communication between proxy and patient (not due to patient's condition)   | 1    | 1   | 2   | 1   |       |     | 2    |
| States that the question is difficult to answer because of overlap with pain, anxiety and depression  |      |     |     |     |       |     |      |
| Unable to give a response (e.g. too distressed or makes reference to the fact that they do not want to think about it)  |      |     |     | 1   |       |     |      |
| Inadequate response (e.g. response does not answer the question, talks about another symptom)   | 1    |     |     | 2   |       |     | 1    |
| Expresses uncertainty (e.g. words such as "probably" or "doubt", or asks for clarification)   | 1    |     | 2   | 3   |       | 2   | 1    |
| Response changed during reply (e.g. says "severe" then changes to "significant", or says "no" to symptom then changes to "yes" during response)   |      |     | 2   | 2   |       |     |      |
| <b>Symptom frequency judgments</b>  |      |     |     |     |       |     |      |
| Symptom constant/almost constant (no frequency estimation)  |      |     | 8   |     | 1     |     | 8    |
| Episode enumeration<br><i>Here the proxy will recall particular episodes within the period of assessment. Their judgment will be based on a count each of the episodes (e.g. "He had pain on the Wednesday and on the Friday in the week before he died")</i>   | 1    |     | 4   |     |       |     | 4    |
| Rate-based estimation<br><i>Here the proxy will base their judgment on the regularity of the occurrence of the symptom (e.g. "She normally got the pain every night")</i>   |      |     | 7   |     |       |     | 7    |
| Enumeration estimation<br><i>Here the proxy will base their judgment on the variable number of occurrences of the symptom. This would include strategies such as averaging over the assessment period or other strategies excluding rated-based estimation and counting/recall episodes (e.g. "It wasn't as often at the beginning but it gradually became worse so over that month I would say often")</i> |      |     | 6   |     |       |     | 6    |

Pain coding

| Code description  | Week |     |     |     | Month |     |      |
|---|------|-----|-----|-----|-------|-----|------|
|   | 1b.  | 1c. | 2b. | 3b. | 8b.   | 9b. | 10b. |
| <b>Comparative judgements</b>   |      |     |     |     |       |     |      |
| <i>Anchoring</i>  |      |     |     |     |       |     |      |
| Anchor on norm what would be expected (e.g. "I would think most people would be depressed if they knew they were going to die")   | 2    |     |     |     | 4     | 1   |      |
| Anchor on self or another (e.g. "I would be screaming with agony if it were me")  |      |     |     |     | 1     |     | 1    |
| General characteristic/trait of the patient (e.g. "He had always been a positive person")   | 2    | 2   | 2   | 6   | 7     |     | 6    |
| Anchor on what they were like previously during the illness (e.g. "he was worse in the last week than in the month")  | 5    | 1   | 2   | 3   | 11    | 2   | 4    |
| Anchor on another response option and adjust (e.g. "It wasn't severe it wasn't as bad as that so I would say significant" or "It was more often than sometimes so I would say often") |      |     | 1   | 3   | 1     |     | 1    |
| <i>Intensity rating questions only</i>  |      |     |     |     |       |     |      |
| Variable symptom severity with rating based on the most frequent rating   |      |     |     | 1   |       |     | 3    |
| Variable symptom severity with rating based on the most severe rating   |      |     |     | 5   |       |     | 6    |
| Variable symptom severity with no explanation for choice of rating  |      |     |     | 1   |       |     |      |
| Intensity rating based on average of variable pain intensity (e.g. "it was severe in the last week, but the week before it was mild, so I would say overall it was moderate")         |      |     |     | 1   |       |     | 3    |
| Frequency used as an indicator of symptom severity  |      |     |     | 1   |       |     | 2    |
| <b>Basis of symptom judgment</b>  |      |     |     |     |       |     |      |
| Reference to verbal communication between patient and proxy   | 11   | 14  | 6   | 8   | 8     | 4   | 4    |
| Reference to non-verbal communication, behaviours or cues to symptom such as the patient's appearance   | 15   | 13  | 6   | 8   | 14    | 8   | 10   |
| Reference to verbal communication with a third party (e.g. nurses, family members)  |      |     |     |     |       |     | 1    |
| Reference to patient's thoughts and feelings (e.g. frustration, concentration, anger)   | 3    | 2   | 2   | 5   | 7     | 1   | 4    |
| Reference to patient's functioning (e.g. activities of daily living, social functioning)  | 6    | 1   | 3   | 6   | 6     | 3   | 6    |
| Reference to patient's coping/control (e.g. references to control or lack of control, coping)   | 3    | 2   | 2   | 2   | 7     | 5   | 1    |
| Reference to medication/help (e.g. explicit references indicating help with the symptom)  | 13   | 5   | 12  | 9   | 14    | 8   | 8    |
| <b>Event cues</b>   |      |     |     |     |       |     |      |
| Person mentioned (e.g. nurses, family)  | 7    | 1   | 4   | 4   | 7     | 1   | 2    |
| Place mentioned (e.g. hospice, home)  | 4    |     | 4   | 4   | 10    | 2   | 4    |
| Characteristic of event mentioned (e.g. last rites, dressing changes, etc.)   | 5    | 1   | 1   | 1   | 5     | 4   | 1    |



Pain coding

| Code description   | Week |     |     |     | Month |     |      |
|--|------|-----|-----|-----|-------|-----|------|
|  | 1b.  | 1c. | 2b. | 3b. | 8b.   | 9b. | 10b. |
| <b>Reference period</b>  |      |     |     |     |       |     |      |
| Anchor on a significant event (e.g. illness, hospital admission)   | 2    |     |     | 1   | 4     | 1   | 3    |
| Anchor on a date (e.g. wedding, holiday)   |      |     |     |     | 1     |     | 1    |
| Anchor on a season of the year (e.g. summer, spring)   |      |     |     |     |       |     |      |
| Mentions reference period being assessed (e.g. week or month)  | 8    | 4   | 8   | 10  | 17    | 4   | 9    |
| Reference to period outside the specified period (last month or week)  | 3    | 1   | 1   | 1   | 2     |     | 1    |
| <b>Proxies' emotions</b>   |      |     |     |     |       |     |      |
| Proxy's emotional reaction to event mentioned  | 1    | 1   | 1   | 2   | 4     | 1   | 1    |
| <b>Affect-based judgments</b>  |      |     |     |     |       |     |      |
| Explicit inferences about how the patient was feeling or thinking based on proxies own feelings about events | 4    | 5   | 3   | 7   | 6     | 3   | 5    |
| <b>Kappa</b>   | .88  | .91 | .83 | .90 | .87   | .91 | .87  |

Anxiety coding

| Code description  | Week |      |      |      | Month |      |      |
|---|------|------|------|------|-------|------|------|
|   | 11b. | 11c. | 12b. | 13b. | 16b.  | 17b. | 18b. |
| <b>Response difficulties</b>  |      |      |      |      |       |      |      |
| <i>Indicates difficulties proxies might experience when answering a question</i>  |      |      |      |      |       |      |      |
| No verbalisation of strategy  |      |      |      |      |       |      |      |
| No apparent strategy (e.g. guess)   |      |      | 1    |      |       |      | 2    |
| Don't know  | 1    |      | 1    |      |       |      | 2    |
| Incoherent response   |      |      |      |      |       |      |      |
| Explicitly states that the question is difficult to answer no reason given why  | 1    |      |      | 1    |       | 1    |      |
| States that the question is difficult to answer because of patient's level of consciousness (e.g. due to medication or confusion)   | 6    | 1    | 1    |      |       | 2    | 1    |
| States that the question is difficult to answer because of lack of communication between proxy and patient (not due to patient's condition)   | 3    |      |      | 2    | 3     | 2    | 2    |
| States that the question is difficult to answer because of overlap with pain, anxiety and depression  | 1    | 1    |      | 1    |       |      | 1    |
| Unable to give a response (e.g. too distressed or makes reference to the fact that they do not want to think about it)  | 1    |      |      | 1    |       |      |      |
| Inadequate response (e.g. response does not answer the question, talks about another symptom)   |      |      | 1    | 1    | 1     |      | 1    |
| Expresses uncertainty (e.g. words such as "probably" or "doubt", or asks for clarification)   | 2    |      |      | 3    | 1     | 2    | 2    |
| Response changed during reply (e.g. says "severe" then changes to "significant", or says "no" to symptom then changes to "yes" during response)   | 1    |      |      |      | 1     | 1    | 1    |
| <b>Symptom frequency judgments</b>  |      |      |      |      |       |      |      |
| Symptom constant/almost constant (no frequency estimation)  |      |      | 3    |      |       | 2    |      |
| Episode enumeration<br><i>Here the proxy will recall particular episodes within the period of assessment. Their judgment will be based on a count each of the episodes (e.g. "He had pain on the Wednesday and on the Friday in the week before he died")</i>   |      |      | 8    |      |       | 6    |      |
| Rate-based estimation<br><i>Here the proxy will base their judgment on the regularity of the occurrence of the symptom (e.g. "She normally got the pain every night")</i>   |      |      | 1    |      |       | 1    |      |
| Enumeration estimation<br><i>Here the proxy will base their judgment on the variable number of occurrences of the symptom. This would include strategies such as averaging over the assessment period or other strategies excluding rated-based estimation and counting/recall episodes (e.g. "It wasn't as often at the beginning but it gradually became worse so over that month I would say often")</i> |      |      | 4    |      |       | 11   |      |

# Anxiety coding

| Code description  | Week |      |      |      | Month |      |      |
|---|------|------|------|------|-------|------|------|
|   | 11b. | 11c. | 12b. | 13b. | 16b.  | 17b. | 18b. |
| <b>Comparative judgements</b>   |      |      |      |      |       |      |      |
| <i>Anchoring</i>  |      |      |      |      |       |      |      |
| Anchor on norm what would be expected (e.g. "I would think most people would be depressed if they knew they were going to die")   | 1    | 2    |      | 1    | 1     |      | 1    |
| Anchor on self or another (e.g. "I would be screaming with agony if it were me")  | 1    |      |      |      |       |      |      |
| General characteristic/trait of the patient (e.g. "He had always been a positive person")   | 5    | 2    | 1    | 6    | 10    | 1    | 2    |
| Anchor on what they were like previously during the illness (e.g. "he was worse in the last week than in the month")  | 2    |      |      | 1    | 5     | 1    | 5    |
| Anchor on another response option and adjust (e.g. "It wasn't severe it wasn't as bad as that so I would say significant" or "It was more often than sometimes so I would say often") |      |      |      | 2    |       | 2    |      |
| <i>Intensity rating questions only</i>  |      |      |      |      |       |      |      |
| Variable symptom severity with rating based on the most frequent rating   |      |      |      |      |       |      | 2    |
| Variable symptom severity with rating based on the most severe rating   |      |      |      |      |       |      | 1    |
| Variable symptom severity with no explanation for choice of rating  |      |      |      | 2    |       |      | 1    |
| Intensity rating based on average of variable pain intensity (e.g. "it was severe in the last week, but the week before it was mild, so I would say overall it was moderate")         |      |      |      |      |       |      | 1    |
| Frequency used as an indicator of symptom severity  |      |      |      | 3    |       |      | 3    |
| <b>Basis of symptom judgment</b>  |      |      |      |      |       |      |      |
| Reference to verbal communication between patient and proxy   | 15   | 7    | 6    | 4    | 12    | 5    | 4    |
| Reference to non-verbal communication, behaviours or cues to symptom such as the patient's appearance   | 8    | 7    | 3    | 5    | 6     | 2    | 3    |
| Reference to verbal communication with a third party (e.g. nurses, family members)  | 2    | 1    | 1    | 1    | 1     | 3    | 1    |
| Reference to patient's thoughts and feelings (e.g. frustration, concentration, anger)   | 17   | 5    | 4    | 7    | 20    | 4    | 6    |
| Reference to patient's functioning (e.g. activities of daily living, social functioning)  | 2    |      |      | 1    | 8     | 1    | 3    |
| Reference to patient's coping/control (e.g. references to control or lack of control, coping)   | 1    |      |      | 2    | 4     | 1    | 4    |
| Reference to medication/help (e.g. explicit references indicating help with the symptom)  | 4    |      |      | 1    | 8     | 1    | 3    |
| <b>Event cues</b>   |      |      |      |      |       |      |      |
| Person mentioned (e.g. nurses, family)  | 7    |      | 2    | 1    | 6     | 2    | 2    |
| Place mentioned (e.g. hospice, home)  | 7    |      | 1    | 1    | 9     | 3    | 2    |
| Characteristic of event mentioned (e.g last rites, dressing changes, etc.)  | 7    |      | 1    | 2    | 4     | 1    | 4    |

Anxiety coding

| Code description   | Week |      |      |      | Month |      |      |
|--|------|------|------|------|-------|------|------|
|  | 11b. | 11c. | 12b. | 13b. | 16b.  | 17b. | 18b. |
| <b>Reference period</b>  |      |      |      |      |       |      |      |
| Anchor on a significant event (e.g. illness, hospital admission)   | 7    | 1    | 3    | 1    | 2     | 2    | 2    |
| Anchor on a date (e.g. wedding, holiday)   | 1    |      |      |      |       |      |      |
| Anchor on a season of the year (e.g. summer, spring)   |      |      |      |      |       |      |      |
| Mentions reference period being assessed (e.g. week or month)  | 15   |      | 7    | 1    | 13    | 10   | 4    |
| Reference to period outside the specified period (last month or week)  | 3    | 1    | 2    | 1    | 5     | 1    | 2    |
| <b>Proxies' emotions</b>   |      |      |      |      |       |      |      |
| Proxy's emotional reaction to event mentioned  | 10   | 1    | 5    | 1    | 5     | 2    |      |
| <b>Affect-based judgments</b>  |      |      |      |      |       |      |      |
| Explicit inferences about how the patient was feeling or thinking based on proxies own feelings about events | 18   | 2    | 3    | 2    | 14    | 4    | 6    |
| <b>Kappa</b>   | .89  | .86  | .89  | .88  | .92   | .87  | .89  |

Depression coding

| Code description  | Week |      |      |      | Month |      |      |
|---|------|------|------|------|-------|------|------|
|   | 19b. | 19c. | 20b. | 21b. | 24b.  | 25b. | 26b. |
| <b>Response difficulties</b><br><i>Indicates difficulties proxies might experience when answering a question</i>  |      |      |      |      |       |      |      |
| No verbalisation of strategy  |      |      |      |      |       |      |      |
| No apparent strategy (e.g. guess)   |      |      |      |      |       |      | 1    |
| Don't know  | 3    |      | 2    | 3    |       |      |      |
| Incoherent response   |      |      |      |      |       |      |      |
| Explicitly states that the question is difficult to answer no reason given why  |      |      |      | 1    |       | 1    | 2    |
| States that the question is difficult to answer because of patient's level of consciousness (e.g. due to medication or confusion)   | 5    | 1    | 2    | 2    | 1     |      |      |
| States that the question is difficult to answer because of lack of communication between proxy and patient (not due to patient's condition)   | 2    | 2    |      | 1    | 4     | 1    | 1    |
| States that the question is difficult to answer because of overlap with pain, anxiety and depression  | 1    | 1    | 2    | 1    | 2     | 1    |      |
| Unable to give a response (e.g. too distressed or makes reference to the fact that they do not want to think about it)  |      |      |      |      | 1     |      |      |
| Inadequate response (e.g. response does not answer the question, talks about another symptom)   | 1    |      |      | 1    |       | 1    |      |
| Expresses uncertainty (e.g. words such as "probably" or "doubt", or asks for clarification)   | 6    |      |      |      |       |      | 3    |
| Response changed during reply (e.g. says "severe" then changes to "significant", or says "no" to symptom then changes to "yes" during response)   |      |      |      |      |       |      | 1    |
| <b>Symptom frequency judgments</b>  |      |      |      |      |       |      |      |
| Symptom constant/almost constant (no frequency estimation)  |      |      | 5    |      |       | 5    |      |
| Episode enumeration<br><i>Here the proxy will recall particular episodes within the period of assessment. Their judgment will be based on a count each of the episodes (e.g. "He had pain on the Wednesday and on the Friday in the week before he died")</i>   |      |      | 4    |      | 1     | 3    |      |
| Rate-based estimation<br><i>Here the proxy will base their judgment on the regularity of the occurrence of the symptom (e.g. "She normally got the pain every night")</i>   |      |      | 2    |      |       | 1    |      |
| Enumeration estimation<br><i>Here the proxy will base their judgment on the variable number of occurrences of the symptom. This would include strategies such as averaging over the assessment period or other strategies excluding rated-based estimation and counting/recall episodes (e.g. "It wasn't as often at the beginning but it gradually became worse so over that month I would say often")</i> |      |      | 2    |      |       | 6    |      |

## Depression coding

| Code description  | Week |      |      |      | Month |      |      |
|---|------|------|------|------|-------|------|------|
|   | 19b. | 19c. | 20b. | 21b. | 24b.  | 25b. | 26b. |
| <b>Comparative judgements</b>   |      |      |      |      |       |      |      |
| <i>Anchoring</i>  |      |      |      |      |       |      |      |
| Anchor on norm what would be expected (e.g. "I would think most people would be depressed if they knew they were going to die")   | 3    | 1    |      | 3    | 5     |      | 5    |
| Anchor on self or another (e.g. "I would be screaming with agony if it were me")  | 2    | 1    | 2    |      | 2     |      | 1    |
| General characteristic/trait of the patient (e.g. "He had always been a positive person")   | 11   | 4    | 1    | 2    | 11    | 4    | 5    |
| Anchor on what they were like previously during the illness (e.g. "he was worse in the last week than in the month")  | 4    |      |      | 1    | 7     | 3    | 7    |
| Anchor on another response option and adjust (e.g. "It wasn't severe it wasn't as bad as that so I would say significant" or "It was more often than sometimes so I would say often") |      |      |      | 3    |       |      | 1    |
| <i>Intensity rating questions only</i>  |      |      |      |      |       |      |      |
| Variable symptom severity with rating based on the most frequent rating   |      |      |      |      |       |      | 2    |
| Variable symptom severity with rating based on the most severe rating   |      |      |      | 2    |       |      |      |
| Variable symptom severity with no explanation for choice of rating  |      |      |      |      |       |      | 1    |
| Intensity rating based on average of variable pain intensity (e.g. "it was severe in the last week, but the week before it was mild, so I would say overall it was moderate")         |      |      |      |      |       |      | 1    |
| Frequency used as an indicator of symptom severity  |      |      |      | 3    |       |      | 3    |
| <b>Basis of symptom judgment</b>  |      |      |      |      |       |      |      |
| Reference to verbal communication between patient and proxy   | 5    | 3    | 4    | 3    | 11    | 3    | 5    |
| Reference to non-verbal communication, behaviours or cues to symptom such as the patient's appearance   | 8    | 7    | 5    | 5    | 8     | 3    | 4    |
| Reference to verbal communication with a third party (e.g. nurses, family members)  |      |      |      |      |       |      |      |
| Reference to patient's thoughts and feelings (e.g. frustration, concentration, anger)   | 16   | 4    | 5    | 6    | 19    | 7    | 9    |
| Reference to patient's functioning (e.g. activities of daily living, social functioning)  | 8    | 2    |      | 3    | 9     | 2    | 6    |
| Reference to patient's coping/control (e.g. references to control or lack of control, coping)   | 2    | 2    | 1    | 1    | 5     |      | 3    |
| Reference to medication/help (e.g. explicit references indicating help with the symptom)  | 3    | 1    | 2    | 1    | 4     |      |      |
| <b>Event cues</b>   |      |      |      |      |       |      |      |
| Person mentioned (e.g. nurses, family)  | 4    |      | 3    | 1    | 5     | 3    | 1    |
| Place mentioned (e.g. hospice, home)  | 2    |      |      |      | 2     | 3    |      |
| Characteristic of event mentioned (e.g. last rites, dressing changes, etc.)   | 2    |      | 4    | 1    | 2     | 2    | 2    |

Depression coding

| Code description   | Week |      |      |      | Month |      |      |
|--|------|------|------|------|-------|------|------|
|  | 19b. | 19c. | 20b. | 21b. | 24b.  | 25b. | 26b. |
| <b>Reference period</b>  |      |      |      |      |       |      |      |
| Anchor on a significant event (e.g. illness, hospital admission)   | 5    |      | 1    | 1    | 3     | 2    | 1    |
| Anchor on a date (e.g. wedding, holiday)   | 2    |      |      |      | 1     |      |      |
| Anchor on a season of the year (e.g. summer, spring)   |      |      |      |      |       |      |      |
| Mentions reference period being assessed (e.g. week or month)  | 4    |      | 4    | 2    | 6     | 4    | 4    |
| Reference to period outside the specified period (last month or week)  | 5    |      |      | 1    | 6     | 1    |      |
| <b>Proxies' emotions</b>   |      |      |      |      |       |      |      |
| Proxy's emotional reaction to event mentioned  | 2    |      | 2    |      | 6     | 2    | 3    |
| <b>Affect-based judgments</b>  |      |      |      |      |       |      |      |
| Explicit inferences about how the patient was feeling or thinking based on proxies own feelings about events | 15   | 3    | 2    | 7    | 9     | 1    | 4    |
| <b>Kappa</b>   |      |      |      |      |       |      |      |
|  | .89  | .86  | .89  | .88  | .92   | .87  | .89  |

**Appendix O**

Analysis of differences in VOICES ratings  
between those with high and low levels of self-  
report caregiver strain



**Analysis of differences in VOICES ratings between those with high and low levels of self-reported anxiety.**

|                              | SAI High |           | SAI Low  |           | Mann-Whitney |          | (K-S)    |
|------------------------------|----------|-----------|----------|-----------|--------------|----------|----------|
|                              | <i>N</i> | Mean Rank | <i>N</i> | Mean Rank | <i>z</i>     | <i>p</i> | <i>p</i> |
| Pain frequency (week)        | 18       | 14.94     | 12       | 16.33     | -.449        | .692     | .948     |
| Pain severity (week)         | 18       | 15.19     | 12       | 15.96     | -.243        | .819     | .988     |
| Pain frequency (month)       | 18       | 15.83     | 12       | 15.00     | -.264        | .819     | 1.000    |
| Pain severity (month)        | 18       | 15.92     | 12       | 14.88     | -.327        | .755     | 1.000    |
| Anxiety frequency (week)     | 18       | 18.64     | 12       | 10.79     | -2.48        | .020     | .081     |
| Anxiety severity (week)      | 18       | 18.25     | 12       | 11.38     | -2.18        | .040     | .081     |
| Anxiety frequency (month)    | 18       | 18.31     | 12       | 11.29     | -2.22        | .030     | .226     |
| Anxiety severity (month)     | 18       | 17.81     | 12       | 12.04     | -1.812       | .079     | .512     |
| Depression frequency (week)  | 18       | 18.06     | 12       | 11.07     | -2.07        | .050     | .120     |
| Depression severity (week)   | 18       | 18.28     | 12       | 11.33     | -2.22        | .040     | .120     |
| Depression frequency (month) | 18       | 17.36     | 12       | 12.71     | -1.486       | .158     | .635     |
| Depression severity (month)  | 18       | 17.71     | 12       | 13.00     | -1.330       | .215     | .759     |

$\alpha = .05.$

**Analysis of differences in VOICES ratings between those with high and low levels of self-reported caregiver strain.**

|                              | SAI High |           | SAI Low  |           | Mann-Whitney |          | (K-S)    |
|------------------------------|----------|-----------|----------|-----------|--------------|----------|----------|
|                              | <i>N</i> | Mean Rank | <i>N</i> | Mean Rank | <i>z</i>     | <i>p</i> | <i>p</i> |
| Pain frequency (week)        | 15       | 16.07     | 15       | 14.93     | -.374        | .740     | .925     |
| Pain severity (week)         | 15       | 18.50     | 15       | 12.50     | -1.947       | .610     | .181     |
| Pain frequency (month)       | 15       | 18.90     | 15       | 12.10     | -2.20        | .030     | .181     |
| Pain severity (month)        | 15       | 18.40     | 15       | 12.60     | -1.861       | .074     | .375     |
| Anxiety frequency (week)     | 15       | 15.53     | 15       | 15.47     | -.021        | 1.000    | 1.000    |
| Anxiety severity (week)      | 15       | 15.27     | 15       | 15.73     | -.151        | .902     | 1.000    |
| Anxiety frequency (month)    | 15       | 16.70     | 15       | 14.30     | -.774        | .461     | .660     |
| Anxiety severity (month)     | 15       | 17.03     | 15       | 13.97     | -.984        | .345     | .660     |
| Depression frequency (week)  | 15       | 14.50     | 15       | 16.50     | -.660        | .539     | .999     |
| Depression severity (week)   | 15       | 14.70     | 15       | 16.30     | -.523        | .624     | .999     |
| Depression frequency (month) | 15       | 16.17     | 15       | 14.83     | -.435        | .683     | .925     |
| Depression severity (month)  | 15       | 16.73     | 15       | 14.27     | -.803        | .461     | .925     |

$\alpha = .05.$

## **Appendix P**

Analysis of differences between proxies who  
were followed-up and those who were not

|                           | Non follow-up | Follow-up | Statistical analysis        | Significance            |
|---------------------------|---------------|-----------|-----------------------------|-------------------------|
| <i>Patient age</i>        |               |           |                             |                         |
| Mean                      | 17.44         | 12.96     | Mann-Whitney (two-tailed)   | $z = -1.39, p = .17$    |
| <i>Proxy age</i>          |               |           |                             |                         |
| Mean                      | 17.55         | 13.35     | Mann-Whitney (two-tailed)   | $z = -1.17, p = .24$    |
| <i>Patient sex</i>        |               |           |                             |                         |
| Male                      | $N=9$         | $N=6$     | Chi-square                  | $\chi^2 = .14, p = .73$ |
| Female                    | $N=8$         | $N=7$     |                             |                         |
| <i>Proxy sex</i>          |               |           |                             |                         |
| Male                      | $N=4$         | $N=7$     | Fisher's exact (two-tailed) | $p = .13$               |
| Female                    | $N=13$        | $N=6$     |                             |                         |
| <i>Patient SES</i>        |               |           |                             |                         |
| Class I                   | $N=2$         | $N=2$     | Chi-square                  | $\chi^2 = .07, p = .79$ |
| Class II                  | $N=2$         | $N=1$     |                             |                         |
| Class III                 | $N=3$         | $N=3$     |                             |                         |
| Class IV                  | $N=6$         | $N=4$     |                             |                         |
| Class V                   | $N=0$         | $N=3$     |                             |                         |
| Class VI                  | $N=4$         | $N=0$     |                             |                         |
| <i>Proxy SES</i>          |               |           |                             |                         |
| Class I                   | $N=1$         | $N=3$     | Chi-square                  | $\chi^2 = 22, p = .64$  |
| Class II                  | $N=3$         | $N=2$     |                             |                         |
| Class III                 | $N=5$         | $N=3$     |                             |                         |
| Class IV                  | $N=3$         | $N=2$     |                             |                         |
| Class V                   | $N=2$         | $N=3$     |                             |                         |
| Class VI                  | $N=3$         | $N=0$     |                             |                         |
| <i>Place of death</i>     |               |           |                             |                         |
| Home                      | $N=6$         | $N=3$     | Fisher's exact (two-tailed) | $p = .69$               |
| Hospice                   | $N=7$         | $N=8$     |                             |                         |
| Hospital                  | $N=4$         | $N=1$     |                             |                         |
| Nursing home              | $N=0$         | $N=1$     |                             |                         |
| <i>Residence</i>          |               |           |                             |                         |
| Residing with patient     | $N=14$        | $N=10$    | Fisher's exact (two-tailed) | $p = 1.00$              |
| Not residing with patient | $N=3$         | $N=3$     |                             |                         |

$\alpha = .05.$